

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 15:59:01 ; Search time 4746.96 Seconds  
(without alignments)  
10471.464 Million cell updates/sec

Title: US-10-010-408-1  
Perfect score: 1708  
Sequence: 1 GAGCGTCTGATCTCCAGAG.....GCCTAGATAAATACACCCAAA 1708

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues  
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_htg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*  
15: em\_ba.\*  
16: em\_fun.\*  
17: em\_hum.\*  
18: em\_in.\*  
19: em\_mu.\*  
20: em\_om.\*  
21: em\_or.\*  
22: em\_ov.\*  
23: em\_pat.\*  
24: em\_ph.\*  
25: em\_pl.\*  
26: em\_ro.\*  
27: em\_sts.\*  
28: em\_un.\*  
29: em\_vi.\*  
30: em\_htg\_hum.\*  
31: em\_htg\_inv.\*  
32: em\_htg\_other.\*  
33: em\_htg\_mus.\*  
34: em\_htg\_pln.\*  
35: em\_htg\_rod.\*  
36: em\_htg\_mam.\*  
37: em\_htg\_vrt.\*  
38: em\_sy.\*  
39: em\_htgo\_hum.\*  
40: em\_htgo\_mus.\*  
41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID | Description        |
|------------|--------|-------------|--------|-------|--------------------|
| 1          | 1621.6 | 94.9        | 1741   | 10    | AF259981 Rattus no |
| 2          | 1278   | 74.8        | 1734   | 6     | AR210324 Sequence  |
| 3          | 1278   | 74.8        | 1734   | 6     | AR210325 Sequence  |
| 4          | 1278   | 74.8        | 1734   | 10    | AF100778 Mus muscu |
| 5          | 1263.8 | 74.0        | 1739   | 10    | AF126063 Mus muscu |
| 6          | 901.2  | 52.8        | 137964 | 2     | AC126895 Rattus no |
| 7          | 881.8  | 51.6        | 226303 | 2     | AC095418 Rattus no |
| 8          | 645    | 37.8        | 61072  | 10    | AL731698 Mouse DNA |
| 9          | 645    | 37.8        | 216757 | 2     | AL669906 Mus muscu |
| 10         | 561.4  | 32.9        | 1266   | 6     | AX076919 Sequence  |
| 11         | 561.4  | 32.9        | 1266   | 6     | AX64186 Sequence   |
| 12         | 561.4  | 32.9        | 1283   | 9     | AF083500 Homo sapi |
| 13         | 561.4  | 32.9        | 1293   | 6     | AR210322 Sequence  |
| 14         | 561.4  | 32.9        | 1293   | 6     | AR210323 Sequence  |
| 15         | 561.4  | 32.9        | 1309   | 9     | AF074604 Homo sapi |
| 16         | 561.4  | 32.9        | 1427   | 9     | AF100780 Homo sapi |
| 17         | 561.4  | 32.9        | 1450   | 9     | BC017782 Homo sapi |
| 18         | 528.8  | 31.0        | 841    | 6     | AR210338 Sequence  |
| 19         | 501.2  | 29.3        | 738    | 6     | AR210337 Sequence  |
| 20         | 256.6  | 15.0        | 137964 | 2     | AC126895 Rattus no |
| 21         | 195.8  | 11.5        | 107260 | 9     | AL139352 Human DNA |
| 22         | 168    | 9.8         | 1585   | 5     | NV1271167          |
| 23         | 163.8  | 9.6         | 2075   | 6     | AR018957 Sequence  |
| 24         | 163.8  | 9.6         | 2075   | 6     | AR118595 Sequence  |
| 25         | 163.8  | 9.6         | 2075   | 6     | AR130891 Sequence  |
| 26         | 163.8  | 9.6         | 2075   | 6     | AX206708 Sequence  |
| 27         | 163.8  | 9.6         | 2075   | 6     | 111636 Sequence 1  |
| 28         | 163.8  | 9.6         | 2075   | 6     | 132210 Sequence 1  |
| 29         | 163.8  | 9.6         | 2075   | 9     | HUMCONGRO          |
| 30         | 163.8  | 9.6         | 2312   | 9     | HSC7GF             |
| 31         | 163.8  | 9.6         | 2998   | 6     | AR119211 Sequence  |
| 32         | 163.8  | 9.6         | 2998   | 6     | AR151276 Sequence  |
| 33         | 161    | 9.4         | 2267   | 6     | AX206706 Sequence  |
| 34         | 161    | 9.4         | 2267   | 10    | MUSFISPI2B         |
| 35         | 161    | 9.4         | 2330   | 10    | BC006783           |
| 36         | 161    | 9.4         | 2334   | 10    | BC006783           |
| 37         | 159.2  | 9.3         | 1496   | 4     | SSU83916           |
| 38         | 158    | 9.3         | 2338   | 6     | E37595             |
| 39         | 158    | 9.3         | 2338   | 10    | AB023068           |
| 40         | 158    | 9.3         | 2345   | 6     | AX402019           |
| 41         | 158    | 9.3         | 2345   | 10    | AF120275           |
| 42         | 156.4  | 9.2         | 2350   | 6     | AR194011           |
| 43         | 156    | 9.1         | 2190   | 5     | AF463517           |
| 44         | 156    | 9.1         | 2288   | 5     | GA298335           |
| 45         | 154.8  | 9.1         | 2350   | 6     | AR201286           |

ALIGNMENTS

RESULT 1  
AF259981  
LOCUS  
DEFINITION Rattus norvegicus CCN family protein COP-1 (Cop-1) mRNA, complete cds.  
ACCESSION AF259981  
VERSION AF259981.1 GI:7739780  
KEYWORDS Rattus norvegicus  
SOURCE Rattus norvegicus  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
REFERENCE 1 (bases 1 to 1741)  
AUTHORS Zhang, R., Averboukh, L., Zhu, W., Zhang, H., Jo, H., Dempsey, P.J.,

Coffey,R.J., Pardee,A.B. and Liang,P.  
Identification of rCop-1, a new member of the CN protein family,  
as a negative regulator for cell transformation  
Mol. Cell. Biol. 18 (10), 6131-6141 (1998)  
98414629  
PUBMED 9742130  
2 (bases 1 to 1741)  
REFERENCE  
AUTHORS Liang,P.  
TITLE Direct Submission  
JOURNAL Submitted (24-APR-2000) Cell Biology, Vanderbilt-Ingram Cancer  
Center, 649 MB II, Nashville, TN 37232, USA  
FEATURES  
source Location/Qualifiers  
1..1741  
/organism="Rattus norvegicus"  
/db\_xref="taxon:10116"  
gene 1..1741  
/gene="Cop-1"  
CDS 262..1014  
/gene="Cop-1"  
/note="secreted protein"  
/codon\_start=1  
/product="CN family protein COP-1"  
/protein\_id="AAF6901.1"  
/db\_xref="GI:7739781"  
/translation="MRGSLIRLIATSLFLLSMVCAQLRTPCTPWTTPQCPQGV  
LVLDGCGCKVCAARLRESCHEHLWCEPSQGLVQPGAGPGHGAALDLDDGDCEV  
NGRRYLDGETFKPNRVLRCDDGGFTCLPLCSEDTPLPSWDCPPKRIQVPGKCP  
WYCDGVTPATORSAAOGLHLSALVTPASADAPWPNWSTANGPCSTCGGLGIATRVSN  
QNRFCOLEIQRLLCLPRCLAAHSNNSAF"  
BASE COUNT 386 a 491 c 480 g 384 t  
ORIGIN  
Query Match 94.9%; Score 1621.6; DB 10; Length 1741;  
Best Local Similarity 98.4%; Pred. No. 0;  
Matches 1680; Conservative 0; Mismatches 24; Indels 4; Gaps 4;  
QY 1 GACGCTTCGTATCTCCAGAGACCCTGGGGTGGGACAGGGGCTTGGCAAGGCTCGAGC 60  
DB 17 GACGCTTCGTATCTCCAGAGACCCTGGGGTGGGACAGGGGCTTGGCAAGGCTCGAG-C 74  
QY 61 GCTGGGAGTGGCTGGGAATGAGGCTTTATTACTGGGAAGCTGAGGAGCTAAGAGGCTC 120  
DB 75 GCTGGGAGTGGCTGGGAATGAGGCTTTATTACTGGGAAGCTGAGGAGCTAAGAGGCTC 134  
QY 121 CTGTACGCTTGCTCTAAAGCTTTAGCACTTGTTGGCTTGGGCTTCACACACTGTCAGA 180  
DB 135 CTGTACGCTTGCTCTAAAGCTTTAGCACTTGTTGGCTTGGGCTTCACACACTGTCAGA 194  
QY 181 CACCTTCGTGGGCTCCAGGCTCAGCTCAGCTTGAAGCTGGCTCCACAGGAC 240  
DB 195 CACCTTCGTGGGCTCCAGGCTCCAGGCTCAGCTTGAAGCTGGCTCCACAGGAC 253  
QY 241 ACAGGTGACATAGGGGAGCCACATGATCTTCTGGCCACTTCTCTCTCTCTCTCTCTCT 300  
DB 254 ACAGGTGACATAGGGGAGCCACATGATCTTCTGGCCACTTCTCTCTCTCTCTCTCTCT 313  
QY 301 TCTCAATGGTGTGCCAGCTGTGCCGACACCCTGTACCTGTCTTGGACACACACCC 360  
DB 314 TCTCAATGGTGTGCCAGCTGTGCCGACACCCTGTACCTGTCTTGGACACACACCC 373  
QY 361 AGTGGCCACAGGGGTACCCCTGGTGGTGGATGGCTGTGGCTGTGAAGTGTGTGAC 420  
DB 374 AGTGGCCACAGGGGTACCCCTGGTGGTGGATGGCTGTGGCTGTGAAGTGTGTGAC 433  
QY 421 GGAGGCTGGGGAGTCTCGACACACCTGCATGTCTGCGACCCAGCCAGGCGCTGGTTT 480  
DB 434 GGAGGCTGACGAGTCTCGGACACCTGCATGTCTGCGACCCAGCCAGGCGCTGGTTT 493  
QY 481 GTACGCTGGGCGAGCCCTGGCGCCATGGGGCTGTGTGTCTTGGATGAGGATGACG 540  
DB 494 GTACGCTGGGCGAGCCCTGGCGCCATGGGGCTGTGTGTCTTGGATGAGGATGACG 553  
QY 541 GTAGCTGTAGGTGAATGGCCGACGCTACCTGGATGGAGAGACCTTTAAACCAATTGCA 600

11 |||||  
554 GTGACTGTGAGTGAATGGCCGACGTAACCTGGATGGAGAGACCTTTAAACCAATTGCA 613  
QY 601 GGGTCTCTGCGCTGTGATGACGGTGGCTTCACTCTGCTGCGCTGTGCAAGTGGAGG 660  
DB 614 GGGTCTCTGCGCTGTGATGACGGTGGCTTCACTCTGCTGCGCTGTGCAAGTGGAGG 673  
QY 661 TCGGCTGCGCTGAGCTGGGACTGCCACGCCCAAGAGAAATACAGGTGCCAGAAAGTCT 720  
DB 674 TGACGCTGCGCTGAGCTGGGACTGCCACGCCCAAGAGAAATACAGGTGCCAGAAAGTCT 733  
QY 721 GCCCGAGTGGGTATGTGACCGAGGAGTGACACCGCGGATCCAGCGCTCCAGGGCGCAAG 780  
DB 734 GCCCGAGTGGGTATGTGACCGAGGAGTGACACCGCGGATCCAGCGCTCCAGGGCGCAAG 793  
QY 781 GACACCAACTTCTGCGCTTGTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840  
DB 794 GACACCAACTTCTGCGCTTGTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 853  
QY 841 GCACAGCTGGGGGCTGCTCAACGACCTGTGGGCTGGGATAGCCACCGAGGTGCTCA 900  
DB 854 GCACAGCTGGGGGCTGCTCAACGACCTGTGGGCTGGGATAGCCACCGAGGTGCTCA 913  
QY 901 ACCAGAACCGATCTGCGCAACTGGAGATCCAAACCGCGCTGTGCTGCGCGAGCCCTGCC 960  
DB 914 ACCAGAACCGATCTGCGCAACTGGAGATCCAAACCGCGCTGTGCTGCGCGAGCCCTGCC 973  
QY 961 TGGCAGCCAGGACGACAGCTCATGGAACAGTCTTCTAAGGCCAAGTGGGATGCGCA 1020  
DB 974 TGGCAGCCAGGACGACAGCTCATGGAACAGTCTTCTAAGGCCAAGTGGGATGCGCA 1033  
QY 1021 TACAGGGCTGCGCATCTCAGCAATAGACCTTAGGACCGAGCCCTGGAGTCTGCTGATAG 1080  
DB 1034 TACAGGGCTGCGCATCTCAGCAATAGACCTTAGGACCGAGCCCTGGAGTCTGCTGATAG 1093  
QY 1081 GCTCTTCTCCAGCTTGGCTGCAATTAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140  
DB 1094 GCTCTTCTCCAGCTTGGCTGCAATTAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1153  
QY 1141 ACTGAGCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200  
DB 1154 ACTGAGCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1213  
QY 1201 GTTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260  
DB 1214 GTTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1273  
QY 1261 GACACCAAAAGACCTTTAAACCTTAGGCTATCTGCGGCAACCTGGCCACCGCTGCTGGG 1320  
DB 1274 GACACCAAAAGACCTTTAAACCTTAGGCTATCTGCGGCAACCTGGCCACCGCTGCTGGG 1333  
QY 1321 GATAAGGTCAATGTTAGGACCAAGACAGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380  
DB 1334 GATAAGGTCAATGTTAGGACCAAGACAGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1393  
QY 1381 ACTTCTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440  
DB 1394 ACTTCTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1453  
QY 1441 TGAGAACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500  
DB 1454 TGAGAACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1513  
QY 1501 GATGAATCACACTGCTCTTAAAGAAATTCCTGAAAGTCCAGGAACCTTGAGCTTGTATTTT 1560  
DB 1514 GATGAATCACACTGCTCTTAAAGAAATTCCTGAAAGTCCAGGAACCTTGAGCTTGTATTTT 1573  
QY 1561 CAGGAATGCACTCTCTTAAAGAAATTCCTGAAAGTCCAGGAACCTTGAGCTTGTATTTT 1620  
DB 1574 CAGGAATGCACTCTCTTAAAGAAATTCCTGAAAGTCCAGGAACCTTGAGCTTGTATTTT 1632  
QY 1621 AGGGCTTCTCTCAGCATGAGAAAGACAGGACAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680



|                       |  |   |                     |
|-----------------------|--|---|---------------------|
| Db                    | 1633   | AGGGCCCTTCTCTTCAGCATGAGAAGACAGAGGGACACAGAGTACTCTCTGCTGAGGA    | 1692                |
| Qy                    | 1681   | CTAGTCTAGCCTAGATAAACCACCAAA                                   | 1708                |
| Db                    | 1693   | CTAGTCTAGCCTAGATAAACCACCAAA                                   | 1720                |
| RESULT 2              |  |   |                     |
| LOCUS                 | AR210324   | 1734 bp   | DNA                 |
| DEFINITION            | Sequence 17 from patent US 6387657.  | linear  | PAT 20-JUN-2002     |
| ACCESSION             | AR210324   |   |                     |
| VERSION               | AR210324.1   | GI:21512525   |                     |
| KEYWORDS              | Unknown.   |   |                     |
| SOURCE                | Unknown.   |   |                     |
| ORGANISM              | Unclassified.  |   |                     |
| REFERENCE             | 1 (bases 1 to 1734)  |   |                     |
| AUTHORS               | Botstein,D.A., Cohen,R.L., Goddard,A.D., Gurney,A.L., Hillan,K.J., Lawrence,D.A., Levine,A.J., Pennica,D., Roy,M. Ann. and Wood,W.I. |   |                     |
| TITLE                 | WISP polypeptides and nucleic acids encoding same  |   |                     |
| JOURNAL               | Patent: US 6387657-A 17 14-MAY-2002;   |   |                     |
| FEATURES              | Location/Qualifiers  |   |                     |
| source                | 1..1734  |   |                     |
| BASE COUNT            | 355 a 491 c 495 g 393 t  |   |                     |
| ORIGIN                | /organism="unknown"  |   |                     |
| Query Match           | 74.8%;   | Score 1278;   | DB 6; Length 1734;  |
| Best Local Similarity | 88.5%;   | Pred. No. 0;  |                     |
| Matches 1520;         | Conservative 0;  | Mismatches 165;   | Indels 33; Gaps 11; |
| Qy                    | 3  | CGCTTCGTATCTCCAGAGACCCCTGGGTGGGACAGAGGGCCCTTGGCAAGGCTGCAGCGC  | 62                  |
| Db                    | 13   | CGCTCTGTATCTCCAGAGACCCCGGCTGGACAGAGGGCCCTTGGCAGGCTGCAGCTGC    | 72                  |
| Qy                    | 63   | TG-GGCAGTGGCTTGAATFAGAGTCTTTTATTACTGGGAAGTGAAGAGGCTCC         | 121                 |
| Db                    | 73   | TGTGGCAGTAGCTTGGGATGGAGTCTTTCTTGTCTGGGAAGTGAAGAGGCTCC         | 132                 |
| Qy                    | 122  | TGTCAG---CTTGCTCTTAAGTCTTAGACATGTGTGGTGGCTTGGCTTCACACACTGTCA  | 178                 |
| Db                    | 133  | TGTCAGGCTCTGTCTTAACACTTTGGCACATGGGTGGCTTGGGCTTCACACACTGTCA    | 192                 |
| Qy                    | 179  | GACACCTTCGTGGTGGCTCCACAGGCCCTCACCTTCAGGTTTGAAGCTGGCTTCCACAAGG | 238                 |
| Db                    | 193  | GACACCTTCGTGGTGGCTCTCGGCC-----TCAGGTTTGAAGCTGGCTTCCACAAGG     | 246                 |
| Qy                    | 239  | ACAGGTGACATAGAGGGGAGCCCACTGATCCATCTTCTGGCCACTTCCTTCCTCTGCCT   | 298                 |
| Db                    | 247  | ACAGGTGACATAGAGGGGCAACCCACTGATCCATCTTCTGGCCATTTCTTCCTCTGCAT   | 306                 |
| Qy                    | 299  | TCCTCAATGCTGTGCCAGCTGTGCCGACACCCTGTACCTGTCTTGGACACACC         | 358                 |
| Db                    | 307  | TCCTCAATGCTGTATCCAGCTGTGCCAGCACCCCTGTGCCCTGTCTTGGACACACC      | 366                 |
| Qy                    | 359  | CCAGTCCCAACAGGGGGTACCCTGGTGGTGGCTGTGGCTGTGGCTGTAAAGTGTGTGC    | 418                 |
| Db                    | 367  | CCAGTCCCAACAGGGGGTACCCTGGTGGTGGCTGTGGCTGTGGCTGTGGCTGTGGC      | 426                 |
| Qy                    | 419  | ACGAGGCTGGGGAGTCTCTGGACCACTGCATGTCTGTGGACCCCAAGCCAGGCGCTGGT   | 478                 |
| Db                    | 427  | ACGAGGCTGGGGAGTCTCTGGACCACTGCATGTCTGCACCCCAAGCCAGGCGCTGGT     | 486                 |
| Qy                    | 479  | TTCTCAGCTGGGCGAGCCCTTGGCGGCATGGGCTGTGTCTCTTGGATGAGGATGA       | 538                 |
| Db                    | 487  | TTGTACGCTGGGCGAGGCCCGAGTGGCGGTGGTGTGTGTGGCTCTTCCGAAGAGGATGA   | 546                 |
| Qy                    | 539  | CGGTAGCTGTAGGTGAATGGCGCAGGTACTGTGATGGAGAGACCTTTAAACCCAATTG    | 598                 |
| Db                    | 547  | CGGGAGCTGTAGGTGAATGGCGCAGGTACTGTGATGGGAGAGACCTTTAAACCCAATTG   | 606                 |
| Qy                    | 599  | CAGGCTCTGTCGCCCTGTGTATGACGGTGGCTTCACCTGTCCTGGCGCTGTGCAGTGA    | 658                 |

Db 1672 CTCTGGAGACTGGCCGGCTGTGGAATAAACACCCAAA 1709

RESULT 3  
AR210325/c  
LOCUS AR210325 1734 bp DNA linear PAT 20-JUN-2002  
DEFINITION Sequence 18 from patent US 6387657.  
ACCESSION AR210325  
VERSION AR210325.1 GI:21512526  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1734)  
AUTHORS Botstein, D.A., Cohen, R.L., Goddard, A.D., Gurney, A.L., Hillan, K.J., Lawrence, D.A., Levine, A.J., Pennica, D., Roy, M. Ann. and Wood, W.I.  
TITLE WISP polypeptides and nucleic acids encoding same  
JOURNAL Patent: US 6387657-A 18-14-MAY-2002;  
FEATURES  
Location/Qualifiers  
1..1734  
/organism="unknown"  
BASE COUNT 393 a 495 c 491 g 355 t  
ORIGIN

Query Match 74.8%; Score 1278; DB 6; Length 1734;  
Best Local Similarity 88.5%; Pred. No. 0;  
Matches 1520; Conservative 0; Mismatches 165; Indels 33; Gaps 11;

QY 3 CCCTTCTGATCTCCAGAGACCTGGGGTGGGACAGGGGCTTGGCAGGCTGCGAGCGC 62  
Db 1722 CCCTCTGATCTCCAGAGACCCCGGCTGGGACAGGGGCTTGGCAGGCTGCGAGCTGC 1663

QY 63 TG-GGCAGTGGCTTGGAAATGGAGTCTTTATTACTGGGAACCTGAGGAGCTAAGAGGCTCC 121  
Db 1662 TGTGGCAGTAGCTGGGATGGAGTCTTTCTTCTGGGACCTGAGAGCTGAGAGGCTCC 1603

QY 122 TGTGAG---CTTGTCTTAAAGTCTTAGCAGTCTGTGGTGGCTTGGGCTTACACACTGTCA 178  
Db 1602 TGTGAGGCTCTCTTAAACTCTTGGGACTTGGGCTGGGCTTGGGCTTACACACTGTCA 1543

QY 179 GACACTTGTGGTGGCTTCCAGGCTTCCAGCTTCCAGCTTGAAGTGGCTTCCAGAGG 238  
Db 1542 GACACTTGTGGTGGCTTCCAGGCTTCCAGCTTCCAGCTTGAAGTGGCTTCCAGAGG 1489

QY 239 ACACGGTGACATGAGGGGAGCCCACTGATCCATCTTCTGGCAGCTTCTTCTCTGCT 298  
Db 1488 ACACGGTGACATGAGGGGAGCCCACTGATCCATCTTCTGGCAGCTTCTTCTCTGCT 1429

QY 299 TCTCTCAATGGTGTGGCAGCTGTGGGACACCCCTGATCTGTCTTGGACACAC 358  
Db 1428 TCTCTCAATGGTGTGGCAGCTGTGGGACACCCCTGATCTGTCTTGGACACAC 1369

QY 359 CCAGTGGCCACAGGGGTACCCCTGGTGGTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 418  
Db 1368 CCAGTGGCCACAGGGGTACCCCTGGTGGTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 1309

QY 419 ACAGGAGCTGGGGAGTCTCTGGACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 478  
Db 1308 ACAGGAGCTGGGGAGTCTCTGGACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1249

QY 479 TTGTGAGCTGGGGAGGCTTGGGGAGTGGGGAGTGGGGAGTGGGGAGTGGGGAGTGGGGAG 538  
Db 1248 TTGTGAGCTGGGGAGGCTTGGGGAGTGGGGAGTGGGGAGTGGGGAGTGGGGAGTGGGGAG 1189

QY 539 CGGTAGCTGAGTGAATGGCCGAGTACCTGATGGATGGATGGATGGATGGATGGATGGAT 598  
Db 1188 CGGTAGCTGAGTGAATGGCCGAGTACCTGATGGATGGATGGATGGATGGATGGATGGATGGAT 1129

QY 599 CAGGTGCTGTGCTGTGATGAGGCTGGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 658  
Db 1128 CAGGTGCTGTGCTGTGATGAGGCTGGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1069

QY 659 TGTGGGCTGGCCAGCTGGGACTGCCACAGGCCCCAAGAGATAACAGGTGCCAGGAAGTG 718

Db 1068 TGTGGGCTGGCCAGCTGGGACTGCCACAGGCCCCAAGAGATAACAGGTGCCAGGAAGTG 1009

QY 719 CTGCCCCAGTGGGTATGTGACAGGAGTGA---CAGCGGGATCCAGCGCTCCAGCGC 775  
Db 1008 CTGCCCCAGTGGGTATGTGACAGGAGTGA---CAGCGGGATCCAGCGCTCCAGCGC 949

QY 776 GCAAGGACACCAACTTTCTGCCCTTGTCTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 835  
Db 948 CCAAGGACACCAACTTTCTGCCCTTGTCTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 889

QY 836 TTGGAGCACAGCTGGGGCCCTGCTCAACACACCTGTGGGCTGGGCTGGGCTGGGCTGGGCT 895  
Db 888 CTGGAGCACAGCTGGGGCCCTGCTCAACACACCTGTGGGCTGGGCTGGGCTGGGCTGGGCT 829

QY 896 GTCCAAACAGTACCACTTCTGCCAACTGGAGATCAACAGCGCGCTGTCTGTGCGCCAGACC 955  
Db 828 ATCCAAACAGTACCACTTCTGCCAACTGGAGATCAACAGCGCGCTGTCTGTGCGCCAGACC 769

QY 956 CTGCTGGCAGCTCAGGAGCCACAGCTCATGGAACAGTGTCTTCTAAGGCCA-ACCTGGGGA 1014  
Db 768 CTGCTGGCAGCTCAGGAGCCAGGCTCATGGAACAGTGTCTTCTAAGGCCA-ACCTGGGGA 709

QY 1015 TCCGATACAGGCGCTGCCATCTCTCAGCAAAATGACCCCTAGGACAGGCGCTGGAGCTGTG 1074  
Db 708 TGTGATACAGGCGCTGCCATCTCTCAGCAAAATGACCCCTAGGACAGGCGCTGGAGCTGTG 649

QY 1075 GTAGATGCTCTCTCCATGCTCTGGCTGCAGTAACTGTCTCTGCTTGGATTCAGTGTCT 1134  
Db 648 GTAGATGCTCTCTCCATGCTCTGGCTGCAGTAACTGTCTCTGCTTGGATTCAGTGTCT 589

QY 1135 AGAGCACTGAGCTGCTCTGCTGTAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1194  
Db 588 AGAGCACTGAGCTGCTCTGCTGTAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 529

QY 1195 TCTCTGG---TTGAGCTTGAATCTGGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1252  
Db 528 TCTCTGGATCTCTGAGCAGGCTTCTGGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 469

QY 1253 ACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1312  
Db 468 ATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 409

QY 1313 GTGCTGGGAGTAAAGCTTAACTGTTAGG-ACCAGACAGCAGATGGCTGAAACTTCCAAATTC 1371  
Db 408 ATGCTGGGAGTAAAGCTTAACTGTTAGG-ACCAGACAGCAGATGGCTGAAACTTCCAAATTC 349

QY 1372 CTTCTTGGACTTCTGTATGTCTTCTCCCAAGATGATGAATGAATGATGATGATGATGATGAT 1431  
Db 348 CTTCTTGGACTTCTGTATGTCTTCTCCCAAGATGATGAATGAATGATGATGATGATGATGAT 290

QY 1432 TCCCTGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1491  
Db 289 TCCCTGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 243

QY 1492 ACATGAAGAGAT-GAATCACACTGTCTTAAAGAAATTCCTGAAAGTCCAGGAACTTGAGC 1550  
Db 242 ACATGAAGAGATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAAT 183

QY 1551 TTTGTATTTTTCAGGAATGCAATCTCTTAAAGCACTGCGCAAAACAGAGAGGCTCCACACT 1610  
Db 182 TTTGTATTTTGAATAAATACATCTCTTAAATGCTCACAAGCAAG-AGGCTCCACACTT 124

QY 1611 CTGGCAGGCGGCTTCTCTCAGCATGAGAAGACACAGGAGACACAGGAGTACTCTC 1670  
Db 123 CTGGCAGGCGGCTTCTCTCAGCATGAGAAGACACAGGAGTACTCTC 64

QY 1671 CTCTGAGGAGTACTCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCT 1708  
Db 63 CTCTGAGGAGTACTCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCT 26

|                       |   |  |                 |                     |
|-----------------------|---|--|-----------------|---------------------|
| AF100778              | 1734 bp   | mRNA   | linear          | ROD 17-DEC-1998     |
| LOCUS                 | Mus musculus connective tissue growth factor related protein WISP-2   |  |                 |                     |
| DEFINITION            | (Wisp2) mRNA, complete cds.   |  |                 |                     |
| ACCESSION             | AF100778  |  |                 |                     |
| VERSION               | AF100778.1 GI:4028578   |  |                 |                     |
| KEYWORDS              | Mus musculus.   |  |                 |                     |
| SOURCE                | Mus musculus  |  |                 |                     |
| ORGANISM              | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  |  |                 |                     |
| REFERENCE             | 1 (bases 1 to 1734)   |  |                 |                     |
| AUTHORS               | Pennica, D., Swanson, T.A., Welsh, J.W., Roy, M.A., Lawrence, D.A., Lee, J., Brush, J., Taneyhill, L.A., Deuel, B., Lew, M., Watanabe, C., Cohen, R.L., Melhem, M.F., Finley, G.G., Quirke, P., Goddard, A.D., Hillan, K.J., Gurney, A.L., Botstein, D., and Levine, A.J. |  |                 |                     |
| TITLE                 | WISP genes are members of the connective tissue growth factor family that are up-regulated in wnt-1-transformed cells and aberrantly expressed in human colon tumors  |  |                 |                     |
| JOURNAL               | Proc. Natl. Acad. Sci. U.S.A. 95 (25), 14717-14722 (1998)   |  |                 |                     |
| MEDLINE               | 99061933  |  |                 |                     |
| PUBMED                | 9843955   |  |                 |                     |
| REFERENCE             | 2 (bases 1 to 1734)   |  |                 |                     |
| AUTHORS               | Pennica, D.   |  |                 |                     |
| Direct Submission     |   |  |                 |                     |
| TITLE                 | Submitted (23-OCT-1998) Molecular Oncology, Genentech, Inc., 1 DNA Way, South San Francisco, CA 94080, USA  |  |                 |                     |
| JOURNAL               | Location/Qualifiers   |  |                 |                     |
| FEATURES              | 1..1734   |  |                 |                     |
| source                | /organism="Mus musculus"  |  |                 |                     |
|                       | /db_xref="taxon:10090"  |  |                 |                     |
|                       | /cell_line="C57MG"  |  |                 |                     |
|                       | /cell_type="epithelial"   |  |                 |                     |
|                       | /tissue_type="mammary"  |  |                 |                     |
|                       | /note="transformed by Wnt-1"  |  |                 |                     |
| gene                  | 1..1734   |  |                 |                     |
| CDS                   | /gene="Wisp2"   |  |                 |                     |
|                       | 257..1012   |  |                 |                     |
|                       | /gene="Wisp2"   |  |                 |                     |
|                       | /codon_start=1  |  |                 |                     |
|                       | /product="connective tissue growth factor related protein WISP-2"   |  |                 |                     |
|                       | /protein_id="AAC96320.1"  |  |                 |                     |
|                       | /db_xref="GI:4028579"   |  |                 |                     |
|                       | /translation="MRGNPLHLIAISFLCILSMVYSQPCAPACBPWPDPQPPGVPLVLDGCCRCVRCARLGECDHLHVCPSQGLVCPGAPSGRGAVCLFEDDGSCEVNGRRYLDGETFKPNCRLVCRDGDGFTCLPLCSFDVRLPSWDCPRPRIOVPGRCPEWVCDQAYMQPAIOPSSAQGHLSALVTPASADGPCPNWSTAWGPCSTTCGLGATRVSNONRFQLEIQRILSLRPLASRSHGSWNSAF"                 |  |                 |                     |
| BASE COUNT            | 355 a   | 491 c  | 495 g           | 393 t               |
| ORIGIN                |   |  |                 |                     |
| Query Match           | 74.8%   | Score 1278;  | DB 10;          | Length 1734;        |
| Best Local Similarity | 88.5%;  | Pred. No. 0;   |                 |                     |
| Matches 1520;         | Conservative  | 0;   | Mismatches 165; | Indels 33; Gaps 11; |
| Qy                    | 3   | CGCTCTGATCTCCAGAGGACCCCTGGGGTGGGACAGGGGCGCTTGGCAAGGCTGCAGCGCG 62   |                 |                     |
| Db                    | 13  | CGCTCTGATCTCCAGAGGACCCCGGGCTGGACAGGGGCGCTTGGCGAGGCTGCAGCTGC 72     |                 |                     |
| Qy                    | 63  | TG-GCGAGTGGTGGGAATGAGTCTTTATTACTGGGAAGTGGAGGCTGAAGAGGCTCC 121      |                 |                     |
| Db                    | 73  | TGTGGCAGTAGTGGGATGGAGGTCTTCTTGTGGGAACTGAGAGGCTGAGAGGCTCC 132       |                 |                     |
| Qy                    | 122   | TGTGAG--CTTGCTTAAGCTTAGCACTTGTGGTGGCTTGGCTTCACACTCTCA 178          |                 |                     |
| Db                    | 133   | TGTCAGGCTCTGTCTTAAGCTTAGCACTTGTGGGCTTGGGCTTCACACTCTCA 192          |                 |                     |
| Qy                    | 179   | GACACCTTCGTGGTGGCTCCACGCGCTCACTTTCAGGTGTTGAAGCTGGGCTCCACAAAGG 238  |                 |                     |
| Db                    | 193   | GACACCTTCGTGGTGGCTCTCGGCC-----TCAGTTGAAGCTGGGCTCCACAAAGG 246       |                 |                     |
| Qy                    | 239   | ACACGGTGACATGAGGGGACGCCACTGATCCATCTTCTGGCCACTTCTCTCTCTCTCTCT 298   |                 |                     |
| 247                   | Db  | ACACGGTGACATGAGGGGCAACCCACTGATCCATCTTCTGGCCACTTCTCTCTCTCTCTCT 306  |                 |                     |
| 299                   | Qy  | TCCTCAATGTTGTCGCCAGCTGTCGGGACACCCCTGTACCTGTCTTCTTGGACACACCC 358    |                 |                     |
| 307                   | Db  | TCCTCAATGTTGTTATTCACAGCTGTGCCAGCACCCCTGTGCTGTCTTCTTGGACACACCC 366  |                 |                     |
| 359                   | Qy  | CCAGTGCACACAGGGGTACCCCTGTGCTGGATGGCTGTGGCTGTCTTAAAGTGTGTGC 418     |                 |                     |
| 367                   | Db  | CCAGTGCACACAGGGGTACCCCTGTGCTGGATGGCTGTGGCTGTCTTAAAGTGTGTGC 426     |                 |                     |
| 419                   | Qy  | ACGAGGCTGGGGAGTCTTGGACACACCTGTGATGTCGACACCCAGCCAGCCGCTGGGT 478     |                 |                     |
| 427                   | Db  | ACGAGGCTGGGGAGTCTTGGACACACCTGTGATGTCGACACCCAGCCAGCCGCTGGGT 486     |                 |                     |
| 479                   | Qy  | TTGTCACCTGGGGACGCGCTTGGCGCCATGGGGCTGTGCTCTTGGATGAGGATGA 538        |                 |                     |
| 487                   | Db  | TTGTCACCTGGGGACGCGCTTGGCGCCATGGGGCTGTGCTCTTGGATGAGGATGA 546        |                 |                     |
| 539                   | Qy  | CGGTAGTGTGAGGTGAATGGCGCAGGTACCTGGATGGAGAGACCTTTAAACCCAAATTG 598    |                 |                     |
| 547                   | Db  | CGGTAGTGTGAGGTGAATGGCGCAGGTACCTGGATGGAGAGACCTTTAAACCCAAATTG 606    |                 |                     |
| 599                   | Qy  | CAGGTCTCTGTGCGCTGTGATGACGGTGGCTTACCTGCTGCGCTGTGCACTGAGGA 658       |                 |                     |
| 607                   | Db  | CAGGTCTCTGTGCGCTGTGATGACGGTGGCTTACCTGCTGCGCTGTGCACTGAGGA 666       |                 |                     |
| 659                   | Qy  | TGTGCGGCTGCCAGCTGGGACTGCCACACCCCAAGAGAAATACAGGTGCCAGGAAAGTG 718    |                 |                     |
| 667                   | Db  | TGTGCGGCTGCCAGCTGGGACTGCCACACCCCAAGAGAAATACAGGTGCCAGGAAAGTG 726    |                 |                     |
| 719                   | Qy  | CTGCCCCAGTGGGTATGTGACCCAGGAGTGA---CACCCGCGATCCAGGCTCCAGGC 775      |                 |                     |
| 727                   | Db  | CTGCCCCAGTGGGTATGTGACCCAGGAGTGA---CACCCGCGATCCAGGCTCCAGGC 786      |                 |                     |
| 776                   | Qy  | GCAAGGACACAACTTCTGCGCTTGTACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 835    |                 |                     |
| 787                   | Db  | CCAAGGACACAACTTCTGCGCTTGTACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 846    |                 |                     |
| 836                   | Qy  | TTGAGGACACAGCTGGGGCGCTCTCAACACCTGTGGTGGGCTAGGACCCCGAGT 895         |                 |                     |
| 847                   | Db  | CTGGAGCACAGCTGGGGCGCTCTCAACACCTGTGGTGGGCTAGGACCCCGAGT 906          |                 |                     |
| 896                   | Qy  | GTCCAACAGAACCGATTCTGCCAACTGGAGATCCAAACGCGCGCTGTGTCTGCCAGACC 955    |                 |                     |
| 907                   | Db  | ATCCAACAGAACCGATTCTGCCAACTGGAGATCCAAACGCGCGCTGTGTCTGCCAGACC 966    |                 |                     |
| 956                   | Qy  | CTGCCTGGCAGCAGGAGGACACAGCTCATGGAACAGTGTCTTCTAAGGCCA-ACTGGGGA 1014  |                 |                     |
| 967                   | Db  | CTGCCTGGCAGTCCAGGAGGACCGGCTCATGGAACAGTGTCTTCTAAGGCCA-ACTGGGGA 1026 |                 |                     |
| 1015                  | Qy  | TGGGATACAGGCGCTGCCATCTCTCAGCAATGACCTAGGACAGGCGCTGGAGTCTGT 1074     |                 |                     |
| 1027                  | Db  | TGTGGATACAGGCGCTGCCATCTCTCAGCAATGACCTAGGACAGGCGCTGGAGTCTGT 1086    |                 |                     |
| 1075                  | Qy  | GTAGATCTCTTCTCCATGCTTCTGGCTGCACTTAAGTCTGCTGCTGCTGCTGCTGCTGCT 1134  |                 |                     |
| 1087                  | Db  | GTAGATCTCTTCTCCATGCTTCTGGCTGCACTTAAGTCTGCTGCTGCTGCTGCTGCTGCT 1146  |                 |                     |
| 1135                  | Qy  | AGAGCCACTGAGCGATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1194   |                 |                     |
| 1147                  | Db  | AGAGCCCTGAGCGATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1206    |                 |                     |
| 1195                  | Qy  | TCCTCTG--TTTCAGCTGGGAATCTCTGGTCTCTCTGCTCATCTCTCAAAACATCCCTGT 1252  |                 |                     |
| 1207                  | Db  | TCCTCTGATCTGACCCAGGCTCTCTGGGTTCTCTCTGCTAGTCTCTCAAAACTTCCCTGT 1266  |                 |                     |
| 1253                  | Qy  | ACAAAAGGACAAACAAAAGACCTTTAAACCTAGGCTATACTGGGCAACCTGGCCACC 1312     |                 |                     |
| 1267                  | Db  | ATGAAAAGGACAAACAAAAGACCTTTAAAGCTAAGCTGTACTGGCAAGCTTGGCCACC 1326    |                 |                     |
| 1313                  | Qy  | GTGCTGGGGATAGGTCAATGTTAGG-ACONAGACAGATTTGCTGAAACTTCCAATTC 1371     |                 |                     |

```

Db      1327 ATGCTGGGATAGTGACAGTAATAGGTACCGAGCAGCAGATTGCTGTAACATCAGGTC 1386
QY      1372 CCTCTTGACATCTGTATGCTGTGTCCTCCCAAGATGATGAATGAACCTGTAAGTGACT 1431
Db      1387 CCTCTTGACATCTGTATGCTGTG-CCCAAGATATGGGTGACCTGTGTAAGTGCTCT 1445
QY      1432 TCCCTGACCTGAGAACACCCCTGCTGCTGGGAAGTATTCAGGGGACAGAAATCTCTGCA 1491
Db      1446 TTCTGTATCTGAGAACACCCCTGCCCGGCTGGGA-----AGAAATTTCTGGGA 1492
QY      1492 ACATGAAGAGAT-GAATCACACTGTCTTAAGAAATCTCTGAAATCTCAGGACCTGAGC 1550
Db      1493 ACATGAAGAGATGGAATCACACTATCTTAAGAGCGTTTGGCAAGTCCAGGAACCTGACC 1552
QY      1551 TTGTGATTTTCAAGGAATGACATCTCTTAAGCACTCGCAAAACAGGAAGGCTCCACACT 1610
Db      1553 TTGTGATTTTGAANAATACACATCTCTTAATGCTCACAAAGCAAG-AGGCTCCACACTT 1611
QY      1611 CTGGCAGGCCAGGGCTTTCTCTTCAGCATGAGAAGACAGAGGACAGAGTAGTACTCTC 1670
Db      1612 CTGGCAGGCCAGGGCTTTCTCTTCAGCATGAGAGACAGGAACAGTAGAGTACCCTC 1671
QY      1671 CTCTGGGACACTAGTCTAGCCTAGATAAACACCCAAA 1708
Db      1672 CTCTGGGACACTGGCCCGGCTGTGGAATAAACACCCAAA 1709

```

## RESULT 5

```

LOCUS      AF126063
DEFINITION Mus musculus connective tissue growth factor-like protein precursor
            (Ctfl) mRNA, complete cds.
ACCESSION  AF126063
VERSION     AF126063.1 GI:4337059
KEYWORDS   Mus musculus
SOURCE     Mus musculus
           Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

## REFERENCE

```

AUTHORS    Kumar,S., Hand,A.T., Connor,J.R., Dodds,R.A., Ryan,P.J.,
            Trill,J.J., Fisher,S.M., Nuttall,M.E., Lipshutz,D.B., Zou,C.,
            Hwang,S.M., Votta,B.J., James,I.E., Rieman,D.J., Gowen,M. and
            Lee,J.C.
TITLE      Identification and cloning of a connective tissue growth
            factor-like cDNA from human osteoblasts encoding a novel regulator
            of osteoblast functions
JOURNAL    J. Biol. Chem. 274 (24), 17123-17131 (1999)
MEDLINE    99287915
PUBMED     10358067
REFERENCE  2 (bases 1 to 1739)
AUTHORS    Kumar,S. and Zou,C.
TITLE      Direct Submission
JOURNAL    Submitted (04-FEB-1999) Bone & Cartilage Biology, UW 2109,
            SmithKline Beecham, 709 Swedeland Rd., King of Prussia, PA 19406,
            USA

```

## FEATURES

```

source     1..1739
            Location/Qualifiers
            /organism="Mus musculus"
            /db_xref="taxon:10090"
            /tissue_type="lung"
gene       1..1739
            /gene="Ctfl"
            242..997
            /gene="Ctfl"
            /note="similar to the Mus musculus WISP-2 protein encoded
            by the sequence presented in GenBank Accession Number
            AF100778; putative growth factor; CTGF-L; contains IGF
            binding (IGFBP), Von Willebrand Factor type C (VWC) repeat
            and thrombospondin type I (TSPI) domains; member of the
            CCN (CTGF/Cyr61/Nov) family; lacks the fourth
            carboxy terminal (CT) domain present in other members of
            the CCN family"
            CDS

```

```

/codon_start=1
/product="connective tissue growth factor-like protein
precursor"
/orfein_id="AAD18058.1"
/db_xref="GI:4337060"
/translation="MRGNPLHLHLAISPLCILSMVYSOLCPAPACNCPWTPPOCPBPVP
LVJLGGCCRYCARRLGSDHLHVCDFSQLVCPQSPGSRGAVCLIFEDDDSCVEY
NGRYLDGETEKPNCRVLCRDGDFCTCLPCSEDLRPLSWDCRPRRIQVPRGCCPE
WVBOAVNQPAPSSAOGHLSALVTPASADGCPNPNWSTANGPCSTTCGLGIATRV
NONRECOLEIQRRLCSRPCLASRSHGWSNAF"
BASE COUNT 375 a 480 c 489 g 395 t
ORIGIN
Query Match 74.0%; Score 1263.8; DB 10; Length 1739;
Best Local Similarity 88.0%; Pred. No. 0;
Matches 1510; Conservative 0; Mismatches 172; Indels 33; Gaps 11;
QY 6 TTCTGATCTCCAGAGGACCCCTGGGGTGGGACAGGGGCTTGGCAAGGCTGCAGCGCTG- 64
Db 1 TCCTGATCTCCAGAGGACCCCGGGCTGGGACAGGGGCTTGGCGAGGCTGCAGCTGCTG 60
QY 65 GGCAGTGGCTGGGAATGAGGCTCTTTATTACTGGGAACCTGAGGAGCTAAGAGGCTCCTGT 124
Db 61 GGCAGTGGCTGGGATGGAGGCTTTTCTGTGGGAACCTGAGGAGCTGAGAAGCTCCTGT 120
QY 125 CAG-----CTTGCTCTAAAGTCTTAGCACTTGTGTGGCTTGGGCTTTCACACACTGTCAGAC 181
Db 121 CAGGCTCCTGCTCTAAACTCTTGGCACTTGGCGTGGCTTGGGCTTTCACACACTGTCAGAC 180
QY 182 ACCTTCTGTGGTGGCTCCAGGCTCCTACCTTCAGTGTGAAAGCTGGCTGCCAAGGGACA 241
Db 181 ACCTTCTGTGGTGGCTCCTCGGCC-----TCAGGTTTGAAGCTGGCTGCCAAGGGACA 234
QY 242 CGGTGACATGAGGGGCAACCCACTGATCCATCTTCTGGCCACTTCTTCTCTGCTGCTTCT 301
Db 235 CGGTGACATGAGGGGCAACCCACTGATCCATCTTCTGGCCACTTCTTCTCTGCTGCTTCT 294
QY 302 CTCATGCTGTGTGCTGCCAGCTGTGCGGACACCCCTGTACCTGTCTCTTGGACACCCACA 361
Db 295 CTCATGCTGTNTTCCAGCTGTGCGGACACCCCTGTGCTTCTTGGACACCCACA 354
QY 362 GTGCCACAGGGGGTACCCCTGGTGTGATGGCTGTGGCTGTCTGTAAAGTGTGTGACAG 421
Db 355 GTGCCACAGGGGGTACCCCTGGTGTGATGGCTGTGGCTGTCTGTGAGTGTGTGCCGG 414
QY 422 GAGGCTGGGGAGTCTCTGCGACCACTGCTGCTGCGACCCACCCAGCGGCTGCTGTG 481
Db 415 GAGGCTGGGGAGTCTCTGCGACCACTGCTGCTGCGACCCACCCAGCGGCTGCTGTG 474
QY 482 TCAGCTGGGGCAGGGCCCTGCGGCGCATGGGCTGTGTCTCTTGGATGAGGATGAGGG 541
Db 475 TCAGCTGGGGCAGGGCCCGGCGGCTGTGTGCTGTGTGCTTTCGAAGAGGATGAGGG 534
QY 542 TAGCTGTGAGGTGAATGGCCGAGTACTGTGATGGAGAGACCTTTAAACCAATTCGAG 601
Db 535 GAGCTGTGAGGTGAACGGCCGACGAGTACCTGTGATGGGAGACCTTTAAACCAATTCGAG 594
QY 602 GGTCTGTGCCGCTGTGATGAGGTTGGCTTACCTGCTGCGGCTGTGCACTGAGGATGT 661
Db 595 GGTCTGTGCCGCTGTGATGAGGTTGGCTTACCTGCTGCGGCTGTGCACTGAGGATGT 654
QY 662 GCGGCTGCCGAGTGGGACTGCCACCGGCCCAAGAGAAATACAGGTGCCAGGAAGTGTGT 721
Db 655 GCGGCTGCCGAGTGGGACTGCCACCGGCCCAAGAGAAATACAGGTGCCAGGAAGTGTGT 714
QY 722 CCCCAGTGGGTATGTGACACAGGAGTGA---CACCGGGATCCAGCGCTCCACGGCGCA 778
Db 715 CCCCAGTGGGTATGTGACACAGGAGTGA---CACCGGGATCCAGCGCTCCACGGCGCA 774
QY 779 AGGACACCAACTTCTGCGGCTTGTGCTCCTGCTGATGCTGCTGCTGCTGCTGCTGCTG 838
Db 775 AGGACACCAACTTCTGCGGCTTGTGCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 834

```



(see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).

\* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 49 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

1 1056: contig of 1056 bp in length  
 1057 1156: gap of unknown length  
 1157 2335: contig of 1179 bp in length  
 2336 2435: gap of unknown length  
 2436 3442: contig of 1007 bp in length  
 3443 3542: gap of unknown length  
 3543 5081: contig of 1539 bp in length  
 5082 5181: gap of unknown length  
 5182 6424: contig of 1243 bp in length  
 6425 6524: gap of unknown length  
 6525 7814: contig of 1290 bp in length  
 7815 7914: gap of unknown length  
 7915 9153: contig of 1239 bp in length  
 9154 9253: gap of unknown length  
 9254 10445: contig of 1192 bp in length  
 10446 10546: gap of unknown length  
 10547 11771: contig of 1226 bp in length  
 11772 11871: gap of unknown length  
 11872 13240: contig of 1369 bp in length  
 13241 13340: gap of unknown length  
 13341 13559: contig of 1019 bp in length  
 13560 14459: gap of unknown length  
 14460 15651: contig of 1192 bp in length  
 15652 15751: gap of unknown length  
 15752 17494: contig of 1743 bp in length  
 17495 17594: gap of unknown length  
 17595 18679: contig of 1085 bp in length  
 18680 18779: gap of unknown length  
 18780 20681: contig of 1902 bp in length  
 20682 20781: gap of unknown length  
 20782 22118: contig of 1337 bp in length  
 22119 22218: gap of unknown length  
 22219 23578: contig of 1360 bp in length  
 23579 23678: gap of unknown length  
 23679 25423: contig of 1745 bp in length  
 25424 25524: gap of unknown length  
 25524 27808: contig of 2285 bp in length  
 27809 27908: gap of unknown length  
 27909 30272: contig of 2364 bp in length  
 30273 30372: gap of unknown length  
 30373 33091: contig of 2719 bp in length  
 33092 33191: gap of unknown length  
 33192 35777: contig of 2586 bp in length  
 35778 35877: gap of unknown length  
 35878 36908: contig of 1031 bp in length  
 36909 37008: gap of unknown length  
 37009 38930: contig of 1922 bp in length  
 38931 39030: gap of unknown length  
 39031 41210: contig of 2180 bp in length  
 41211 41310: gap of unknown length  
 41311 42937: contig of 1627 bp in length  
 42938 43037: gap of unknown length  
 43038 46120: contig of 3083 bp in length  
 46121 46220: gap of unknown length  
 46221 48217: contig of 1997 bp in length  
 48218 48317: gap of unknown length  
 48318 50433: contig of 2116 bp in length  
 50434 50533: gap of unknown length  
 50534 52880: contig of 2347 bp in length  
 52881 52980: gap of unknown length  
 52981 55194: contig of 2214 bp in length  
 55195 55294: gap of unknown length  
 55295 57482: contig of 2188 bp in length  
 57483 57582: gap of unknown length

57583 60781: contig of 3199 bp in length  
 60782 60881: gap of unknown length  
 60882 62599: contig of 1718 bp in length  
 62600 62699: gap of unknown length  
 62700 65203: contig of 2504 bp in length  
 65204 65303: gap of unknown length  
 65304 66914: contig of 1611 bp in length  
 66915 67014: gap of unknown length  
 67015 70839: contig of 3825 bp in length  
 70840 70939: gap of unknown length  
 70940 75139: contig of 4200 bp in length  
 75140 75239: gap of unknown length  
 75240 78296: contig of 3057 bp in length  
 78297 78396: gap of unknown length  
 78397 83138: contig of 4742 bp in length  
 83139 83238: gap of unknown length  
 83239 88204: contig of 4966 bp in length  
 88205 88304: gap of unknown length  
 88305 92238: contig of 3934 bp in length  
 92239 92338: gap of unknown length  
 92340 97439: contig of 5001 bp in length  
 97440 103534: contig of 6095 bp in length  
 103535 103634: gap of unknown length  
 103635 107080: contig of 3446 bp in length  
 107081 107180: gap of unknown length  
 107181 115201: contig of 8021 bp in length  
 115202 115301: gap of unknown length  
 115302 121237: contig of 5936 bp in length  
 121238 121337: gap of unknown length  
 121338 127910: contig of 6573 bp in length  
 127911 128010: gap of unknown length  
 128011 137964: contig of 9954 bp in length.

FEATURES  
 source

1..137964  
 /organism="Rattus norvegicus"  
 /db\_xref="taxon:10116"  
 /clone="CH230-301E4"

BASE COUNT 31831 a 34027 c 35229 g 31979 t 4898 others  
 ORIGIN

Query Match 52.8%; Score 901.2; DB 2; Length 137964;

Best Local Similarity 98.5%; Pred. No. 1.1e-233;  
 Matches 920; Conservative 0; Mismatches 13; Indels 1; Gaps 1;

Qy 775 CGCAAGGACACCACTTTTCTGCCCTTGCTGCTCAACCACTGTGGGCTGGGCATAGCCACCCGAG 834  
 Db 44061 CCCAGGACACCACTTTTCTGCCCTTGCTGCTCAACCACTGTGGGCTGGGCATAGCCACCCGAG 44120  
 Qy 835 ATGGAGCACAGGCTGGGGCCCCCTGCTCAACCACTGTGGGCTGGGCATAGCCACCCGAG 894  
 Db 44121 ATGGAGCACAGGCTGGGGCCCCCTGCTCAACCACTGTGGGCTGGGCATAGCCACCCGAG 44180  
 Qy 895 TGTCACACCAACCGGATTCGCCAATCGGAGATCCACGCGCGCTGTGTGCCACAC 954  
 Db 44181 TGTCACACCAACCGGATTCGCCAATCGGAGATCCACGCGCGCTGTGTGCCACAC 44240  
 Qy 955 CTGCTGGCAGCCAGGAGCCACAGCTCATGAAACAGTGTCTTAAGCCCAACTGGGGA 1014  
 Db 44241 CTGCTGGCAGCCAGGAGCCACAGCTCATGAAACAGTGTCTTAAGCCCAACTGGGGA 44300  
 Qy 1015 TGGGATACAGGCGCTGCCATCCTCAGCAAAATGACCCCTAGGACACGCGCCCTGGAGTGTGTG 1074  
 Db 44301 TGGGATACAGGCGCTGCCATCCTCAGCAAAATGACCCCTAGGACACGCGCCCTGGAGTGTGTG 44360  
 Qy 1075 GTATAGCTCTTCCATGCTCTTGGGTGCGAGTTAACTGTCTCTGGATTCAGTGTGT 1134  
 Db 44361 GTATAGCTCTTCCATGCTCTTGGGTGCGAGTTAACTGTCTCTGGATTCAGTGTGT 44420  
 Qy 1135 AGAGCCACTGAGCGATCCCTCTCTGTCTGAGGTAGGCGGAGCAGGTGACCAAGTCCAGT 1194  
 Db 44421 AGAGCCACTGAGCGATCCCTCTCTGTCTGAGGTAGGCGGAGCAGGTGACCAAGTCCAGT 44480

```
Oy 1195 TCTCTGTTGAGCTGGAATTCCTGGGTTCTCTGGGTCATATTCCTCAAAACATCCCTGTAC 1254
Dy 44481 TCTCTGTTGAGCTGGAATTCCTGGGTTCTCTGGGTCATATTCCTCGGAACATCCCTGTAC 44540
Oy 1255 AAAAAGCACACCAAAAGACCTTTAAACCTAGGCTATATCTGGGCAAACTGGCCACCGT 1314
Dy 44541 AAAAAGCACACCAAAAGACCTTTAAACCTAGGCTATATCTGGGCAAACTGGCCACCGT 44600
Oy 1315 GCTGGGGAATAGGTCATATTTAGGACACGACGACGATTCGCTGAAATTCCTCAATTCCT 1374
Dy 44601 GCTGGGGAATAGGTC-TTGTAGGACACGACGACGATTCGCTGAAATTCCTCAATTCCT 44659
Oy 1375 TCTTGGACTCTGTATGCTTGTCCCAAGATGATGATGATGATGATGATGATGATGATGAT 1434
Dy 44660 TCTTGGACTCTGTATGCTTGTCCCAAGATGATGATGATGATGATGATGATGATGATGAT 44719
Oy 1435 CTGACCTGAGAACACCTGCTGCTCGGGAAGTATTCAGGGGAGATTCCTGTGAACA 1494
Dy 44720 CTGACCTGAGAACACCTGCTGCTCGGGAAGTATTCAGGGGAGATTCCTGTGAACA 44779
Oy 1495 TGAAGAGATGAATCACACTGCTCTTAAGAAATTCCTGAAAGTCCAGGAACCTTGAGCTTTG 1554
Dy 44780 TGAAGAGATGAATCACACTGCTCTTAAGAAATTCCTGAAAGTCCAGGAACCTTGAGCTTTG 44839
Oy 1555 TATTTTCAGAAATGCACATCTTTAAGCACTCGCAAAACAGGAGGCTCCACACTCTGG 1614
Dy 44840 TATTTTCAGAAATGCACATCTTTAAGCACTCGCAAAACAGGAGGCTCCACACTCTGG 44899
Oy 1615 CAGGCCAGGCTTCTCTTTCAGCATGAGAAAGCAAGGACAGCAGACTCTCTCTCT 1674
Dy 44900 TAGGCCAGGCTTCTCTTTCAGCATGAGAAAGCAAGGACAGCAGACTCTCTCTCT 44959
Oy 1675 GGAGGACTAGTCTAGGCTAGATAAACCACCA 1708
Dy 44960 GGAGGACTAGTCTAGGCTAGATAAACCACCA 44993
```

## RESULT 7

```
AC095418
LOCUS Rattus norvegicus clone CH230-7C10, *** SEQUENCING IN PROGRESS ***
DEFINITION 226303 bp DNA linear HTG 11-JUL-2002
51 unordered pieces.
AC095418
VERSION AC095418.3 GI:21717893
KEYWORDS HTG; HTGS_PHASE1.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
```

## REFERENCE

```
1 (bases 1 to 226303)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaral-Tunney,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
Bouch,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,K., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Dunthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Ehrhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseghe,H.,
```

```
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
Oragunye,N., Orvedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savary,G.,
Scherer,S., Scott,G., Shen,H., Shoshitari,N., Sisson,I.,
Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansley,J., Taylor,C., Taylor,T., Teifrod,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G., and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 226303)
Worley,K.C.
Direct Submission
Submitted (17-SEP-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 226303)
Worley,K.C.
Direct Submission
Submitted (11-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 10, 2002 this sequence version replaced gi:17941885.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GCDP
Center clone name: CH230-7C10
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 185950 bases at least Q40
Consensus quality: 190362 bases at least Q30
Consensus quality: 193076 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank.draft.data.html).
* NOTE: This is a 'working draft' sequence. It currently
consists of 51 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.
* 1 1051: contig of 1051 bp in length
* 1052 1151: gap of unknown length
* 1152 2286: contig of 1135 bp in length
* 2287 2386: gap of unknown length
* 2387 3433: contig of 1047 bp in length
* 3434 3533: gap of unknown length
* 3534 4614: contig of 1081 bp in length
* 4615 4715: gap of unknown length
* 4715 6581: contig of 1867 bp in length
* 6582 7773: gap of unknown length
* 7774 7873: contig of 1092 bp in length
* 7874 9530: contig of 1657 bp in length
* 9531 9630: gap of unknown length
```







QY 1492 ACATGAGAGATGAATCACACTCTCTTAAGAAATTCCTCAAGTCCAGGAACCTTGAGCT 1551  
 |||||  
 Db 112970 ACATGAGAGATGAATCACACTCTCTTAAGAAATTCCTCAAGTCCAGGAACCTTGAGCT 113029  
 |||||  
 QY 1552 TTGTATTTTTCAGGAATGCACATCTCTTAAGCACTCGCAAAACAGGAAGGCTCCACACCTC 1611  
 |||||  
 Db 113030 TTGTATTTTTCAGGAATGCACATCTCTTAAGCACTCGCAAAACAGGAAGGCTCCACACCTC 113089  
 |||||  
 QY 1612 TGGCAGCCAGGCGCTTCTCTTTCAGCATGAGAAACAGGAAGGACAGCAGAGTACTCTCC 1671  
 |||||  
 Db 113090 TGGTAGCCAGGCGCTTCTCTTTCAGCATGAGAAACAGGAAGGACAGCAGAGTACTCTCC 113149  
 |||||  
 QY 1672 TCTGGAGGACTACTCTAGCTAGCAATAAACACCCAAA 1708  
 |||||  
 Db 113150 TCTGGAGGACTAGTCCAGGCTAGCAATAAACACCCAAA 113186  
 |||||

## RESULT 8

AL731698

LOCUS AL731698 61072 bp DNA linear ROD 24-MAY-2002  
 DEFINITION Mouse DNA sequence from clone RP23-161B3 on chromosome 2, complete sequence.

ACCESSION

AL731698

VERSION

AL731698.10

KEYWORDS

HTG.

SOURCE

house mouse.

ORGANISM

Mus musculus

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Submitted (24-MAY-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk  
 On May 25, 2002 this sequence version replaced g1:21213601.  
 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.  
 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at  
[http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep)  
 from the RPI-23 Mouse PAC Library  
 constructed by the group of Pieter de Jong.  
 For further details see <http://www.chori.org/bacpac/home.htm>  
 VECTOR: pBACe3.6.

## FEATURES

source

1. .61072 Location/Qualifiers  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /chromosome="2"  
 /clone="RP23-161B3"  
 /clone\_lib="RPCI-23"  
 BASE COUNT 15405 a 15368 c 15033 g 15266 t  
 ORIGIN

Query Match

Best Local Similarity

Matches

804; Conservative

0; Mismatches

115; Indels

20; Gaps

7;

QY 775

CGCAAGGACACCAACTTCTGCGCTGTCTACTCTGCTGATGCTCTTGTCCTTGTC

CGCA

15998 CCCAGGACACCAACATTTCTGCGCTTGTCACTCTCTGATCTGCCGATGGCCCTTGCCAA 17057  
 |||||  
 QY 835 ATTGGAGCAGAGCTGGGGCCCTGCTCAACCACTCTGCTGGCTGGGCTAGACCCACCGAG 894  
 |||||  
 Db 17058 ACTGGAGCAGAGCTGGGGCCCTGCTCAACCACTCTGCTGGCTGGGCTAGACCCACCGAG 17117  
 |||||  
 QY 895 TGTCACACAGAACCGATTCTGCCAATCGAGATCCAAACCCGCCCTGTGTCTGCCACAG 954  
 |||||  
 Db 17118 TATCCAAACCAAGAACCGATTCTGCCAATCGAGATCCAAACCCGCCCTGTGTCTGCCAG 17177  
 |||||  
 QY 955 CTTGCTGGCAGCAGGAGCAGCTCATGGAACAGTCTTCTTAAGGCCA-ACCTGGGG 1013  
 |||||  
 Db 17178 CTTGCTGGCAGCAGGAGCAGCTCATGGAACAGTCTTCTTAAGGCCA-ACCTGGGG 17237  
 |||||  
 QY 1014 ATCGGATACAGGGCTGCCATCTCAGCAAAATGACCTTAGGACCAAGCCCTGGAGCTGT 1073  
 |||||  
 Db 17238 ATGTGGATACAGGGCTGCCATCTCAGCAAAATGACCTTAGGACCAAGCCCTGGAGCTGT 17297  
 |||||  
 QY 1074 GGTAGATGCTTCTTCATGCTCTTGGCTGACGTTAACTGCTCTGCTTGGATTCACATGTG 1133  
 |||||  
 Db 17298 GGTAGATGCTTCTTCATGCTCTTGGCTGACGTTAACTGCTCTGCTTGGATTCACATGTG 17357  
 |||||  
 QY 1134 TAGAGCACTGAGCGATCCCTGCTCTGAGGTAGGGGAGCAGGTGACCAAGCTCCAG 1193  
 |||||  
 Db 17358 CAGAGCTCTGAGCGATCCCTGCTCTGAGGTAGGGGAGCAGGTGACCAAGCTCCAT 17417  
 |||||  
 QY 1194 TTCTCTGG--TTGAGCTGGAATTTCTGGGTTCTCTGGCTTCATCTCTCAAAACATCCCTG 1251  
 |||||  
 Db 17418 TTCTCTGGATTTCTGACCCAGGCTTCTGGGTTCTCTGGCTAGTTCTCTCAAAACATCCCTG 17477  
 |||||  
 QY 1252 TACAAAAGGACAAACCAAAAGACCTTTAAACCTAGGCTATATCTGGGCAAAACCTGGCCAC 1311  
 |||||  
 Db 17478 TATGAAAAGGACAAACCAAAAGGACCTTTAAAGCTAAAGCTGCTACTGGGCAAGCCTGGCCAC 17537  
 |||||  
 QY 1312 CGTCTGGGGATAAGTCAATGTTAGG-ACCAGACAGCAGATGCTCTGAAACTTCCCAATT 1370  
 |||||  
 Db 17538 CATGCTGGGATAGTACAGTAATAGTACAGGACAGATGCTCTGAAACATCCAGGT 17597  
 |||||  
 QY 1371 CCCTTCTGGACTTCTGTATGCTTGTCCCAAGATGATGAATGAATTCGTAAGTGPACC 1430  
 |||||  
 Db 17598 CCCTTCTGGACTTCTGTATGCTTGT-CCCAAGATATATGGGTGACCTTCTTAAGTGTGCC 17656  
 |||||  
 QY 1431 TTCCCTGACCTGAGACACACCTGCTGCTGGGAAGTATTCAGGGGCGCAGAAATTCCTGTG 1490  
 |||||  
 Db 17657 TTTCCTGATCTGAGAACACCTGCCCGCTGGGA-----AGAAATTTCTCTGGG 17703  
 |||||  
 QY 1491 AACATGAAGAGAT-GAATCACACTGTCTTAAAGAAATTCCTGAAAGTCCAGGAACCTTGAG 1549  
 |||||  
 Db 17704 AACATGAAGAGATGGAATGGAATCACACTATTTTAAAGAGGCTTTGCCAAGTCCAGGAACCTTGAC 17763  
 |||||  
 QY 1550 CTTTGTATTTCAGGAATGCACATCTCTTAAGCACTCGCAAAACAGGAAGGCTCCACACC 1609  
 |||||  
 Db 17764 CTTTGTATTTCAGGAATGCACATCTCTTAAGCACTCGCAAAACAGGAAGGCTCCACACT 17822  
 |||||  
 QY 1610 TCTGGAGCCAGGCGCTTCTCTTCTCAGCATGAGAAAGACAGGACAGCAGAGTACTCT 1669  
 |||||  
 Db 17823 TCTGGAGCCAGGCGCTTCTCTTCTCAGCATGAGAAAGACAGGACAGTACTCT 17882  
 |||||  
 QY 1670 CTTCTGGAGGACTAGTCTAGCCTAGAAATAAACACCCAAA 1708  
 |||||  
 Db 17883 CTTCTGGAGGACTAGCCTAGCCTAGAAATAAACACCCAAA 17921  
 |||||

## RESULT 9

AL669906

LOCUS

DEFINITION

Mus musculus chromosome 2 clone RP23-217C2, \*\*\* SEQUENCING IN

PROGRESS \*\*\*

AL669906

ACCESSION

VERSION

AL669906.5

KEYWORDS

HTG: HTGS PHASE1; HTGS\_CANCELLED.

SOURCE

Mus musculus.

Query Match

Best Local Similarity

Matches

804; Conservative

0; Mismatches

115; Indels

20; Gaps

7;

QY 775

CGCAAGGACACCAACTTCTGCGCTGTCTACTCTGCTGATGCTCTTGTCCTTGTC

CGCA

## ORGANISM

Mus musculus

## REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

## AUTHORS

Sims,S.

## JOURNAL

Direct Submission  
Submitted (23-JUL-2002) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
humquery@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk  
On Jul 25, 2002 this sequence version replaced gi:18181793.

## COMMENT

----- Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: <http://www.sanger.ac.uk>  
Contact: [humquery@sanger.ac.uk](mailto:humquery@sanger.ac.uk)  
----- Project Information  
Center project name: BM217C2  
----- Summary Statistics  
Assembly program: XGAP4; version 4.5  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Consensus quality: 211865 bases at least Q40  
Consensus quality: 213403 bases at least Q30  
Consensus quality: 214139 bases at least Q20  
Insert size: 214757; sum-of-contigs  
Insert size: 234243; 1.8% error; agarose-fp  
Quality coverage: 6.29x in Q20 bases; sum-of-contigs Quality  
coverage: 5.90x in Q20 bases; agarose-fp  
-----

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 21 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 4000: contig of 4000 bp in length  
\* 4001 4100: gap of 100 bp  
\* 4101 15267: contig of 11167 bp in length  
\* 15268 15367: gap of 100 bp  
\* 15368 34970: contig of 19603 bp in length  
\* 34971 35070: gap of 100 bp  
\* 35071 39671: contig of 4601 bp in length  
\* 39672 39771: gap of 100 bp  
\* 39772 46360: contig of 6589 bp in length  
\* 46361 46460: gap of 100 bp  
\* 46461 49809: contig of 3349 bp in length  
\* 49810 49909: gap of 100 bp  
\* 49910 54509: contig of 4600 bp in length  
\* 54510 54609: gap of 100 bp  
\* 54610 65989: contig of 11380 bp in length  
\* 65990 66089: gap of 100 bp  
\* 66090 84635: contig of 18546 bp in length  
\* 84636 84735: gap of 100 bp  
\* 84736 87399: contig of 2864 bp in length  
\* 87400 87499: gap of 100 bp  
\* 87500 98601: contig of 11102 bp in length  
\* 98602 98701: gap of 100 bp  
\* 98702 105522: contig of 6821 bp in length  
\* 105523 105622: gap of 100 bp  
\* 105623 119773: contig of 14151 bp in length  
\* 119774 119873: gap of 100 bp  
\* 119874 134552: contig of 14679 bp in length  
\* 134553 134652: gap of 100 bp  
\* 134653 141202: contig of 6550 bp in length  
\* 141203 141302: gap of 100 bp  
\* 141303 145808: contig of 4506 bp in length  
\* 145809 145908: gap of 100 bp  
\* 145909 151572: contig of 5664 bp in length  
\* 151573 151672: gap of 100 bp  
\* 151673 169686: contig of 18014 bp in length  
\* 169687 169786: gap of 100 bp  
\* 169787 200095: contig of 30309 bp in length

FEATURES  
source

\* 200096 200195: gap of 100 bp  
\* 200196 213877: contig of 13682 bp in length  
\* 213878 213977: gap of 100 bp  
\* 213978 216757: contig of 2780 bp in length.

Location/Qualifiers  
1..216757  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/chromosome="2"  
/clone\_RP23-217C2"  
/clone\_lib="RPCI-23"  
1..4000  
/note="assembly\_fragment:03429  
clone\_end:SP6  
vector\_side:left"  
4101..15267  
/note="assembly\_fragment:00814  
fragment\_chain:1"  
15368..34970  
/note="assembly\_fragment:00812  
fragment\_chain:1"  
35071..39671  
/note="assembly\_fragment:03544  
fragment\_chain:1"  
39772..46360  
/note="assembly\_fragment:01438  
fragment\_chain:1"  
46461..49809  
/note="assembly\_fragment:03410  
fragment\_chain:2"  
49910..54509  
/note="assembly\_fragment:00811  
fragment\_chain:2"  
54610..65989  
/note="assembly\_fragment:02217  
fragment\_chain:2"  
66090..84635  
/note="assembly\_fragment:03529  
fragment\_chain:2"  
84736..87399  
/note="assembly\_fragment:01019  
fragment\_chain:3"  
87500..98601  
/note="assembly\_fragment:01170  
fragment\_chain:3"  
98702..105522  
/note="assembly\_fragment:02829  
fragment\_chain:3"  
105623..119773  
/note="assembly\_fragment:01476  
fragment\_chain:4"  
119874..134552  
/note="assembly\_fragment:00505  
fragment\_chain:4"  
134653..141202  
/note="assembly\_fragment:01951  
fragment\_chain:4"  
141303..145808  
/note="assembly\_fragment:02637  
fragment\_chain:5"  
145909..151572  
/note="assembly\_fragment:02352  
fragment\_chain:5"  
151673..169686  
/note="assembly\_fragment:02207  
fragment\_chain:5"  
169787..200095  
/note="assembly\_fragment:02849  
fragment\_chain:6"  
200196..213877  
/note="assembly\_fragment:01052  
fragment\_chain:6"  
213978..216757



|            |      |   |      |
|------------|------|---|------|
| QY         | 603  | GTCTCTGGCCGCTGTGATGACGGTGTGCTTACCTGCTGCGCTGTGCAGTGGAGATGTG    | 562  |
| Db         | 364  | ATCCGCTGCGCTGCGAGGACGGCGGCTTACCTGCGTCCGCTGTGCAGCGAGGATGTG     | 423  |
| QY         | 663  | CGGCTGCCAGCTGGGACATGCCACGCCCAAGAGAAATACAGTGTGCCAGGAAAGTCTGCTG | 722  |
| Db         | 424  | CGGCTGCCAGCTGGGACATGCCACGCCCAAGAGGAGGCTGTGCTGCGCAAGTCTGCTG    | 483  |
| QY         | 723  | CCGAGTGGGATGTGACCAAGGAGTGTGACACCGGGCATCCAGCGCTCCACGGCGCAAGGA  | 782  |
| Db         | 484  | CCTGAGTGGGATGTGCGCGCAAGGAGGGGACCTGGGACCCAGGCCCTTCCAGGCCAAGGA  | 543  |
| QY         | 783  | CACCAACTTTGGCCCTTGTCACTCTGCTGTGCTGTGATGCTCTTGTCCAATTTGGAGC    | 842  |
| Db         | 544  | CCCCAGTTTGGGCTTGTCTCTTCCCTGCCGCCCTGGTGTGCCCTGCCAAGTGGAGC      | 603  |
| QY         | 843  | ACAGCCTGGGGCCCTGCTCAACACACTGTGGGCTGGGCATAGCCACCGGAGTGTGCCAAC  | 902  |
| Db         | 604  | ACGGCTGGGACCTGCTCGACCACTGTGGGCTGGGCATGGCCATGCCACCGGGTGTCCAAC  | 663  |
| QY         | 903  | CAGAACCGATTTGCCCACTGGAGATCCAACCGCGCTGTGTGTGCCACAGACCTGCTG     | 962  |
| Db         | 664  | CAGAACCGCTTGTGCGGACTGGAGACCCAGCGCGCTGTGCTGTCCAGGCCCTGCCCA     | 723  |
| QY         | 963  | GCAGCCAGAGCCACAGCTCATGGAACAGTCTTCTTA-AGGCCAAGTGGGATGCGGAT     | 1021 |
| Db         | 724  | CCCTCCAGGGGTGCGAGTCCACAAAACAGTGCCTTAGAGCGCGCTGGAAATGGGAC      | 783  |
| QY         | 1022 | ACAGGCGCTTGCATCTCAGCAAAATCACCTTAGGACAGCGCCCTGGACTGCTGGTAGATG  | 1081 |
| Db         | 784  | ACGGTGTCCACATCCAGCTGGTGGCCCTGTGCTTGGCCCTGGGCTGATGGAAGATG      | 843  |
| QY         | 1082 | CT-CTTCTCTGCTCTTGGCTGTCAGTTAACTGCTGCTTGGATTCACCTGTGTAGAGC     | 1139 |
| Db         | 844  | GTCCGTCGCAAGCCCTTGGCTGTCAGGCAACACTTTAGCTTGGCTGCACCATGCAGAAC   | 902  |
| RESULT 12  |      |   |      |
| AF083500   |      |   |      |
| LOCUS      |      |   |      |
| DEFINITION |      |   |      |
| ACCESSION  |      |   |      |
| VERSION    |      |   |      |
| KEYWORDS   |      |   |      |
| SOURCE     |      |   |      |
| ORGANISM   |      |   |      |
| REFERENCE  |      |   |      |
| AUTHORS    |      |   |      |
| TITLE      |      |   |      |
| JOURNAL    |      |   |      |
| REFERENCE  |      |   |      |
| AUTHORS    |      |   |      |
| TITLE      |      |   |      |
| JOURNAL    |      |   |      |
| FEATURES   |      |   |      |
| source     |      |   |      |
| CDS        |      |   |      |

[illegible]

|  |   |      |   |      |
|--|---|------|---|------|
| D  | b | 783  | ACGGTGTCCACATCCCAGCTGGTGGCCCTGTGCCTTGGGCCCTGGGCTGATGAAGATG    | 842  |
| Q  | y | 1082 | CT-CTTTCTCATCTCTTGCGTCAGATTAACTGTCTCTGCTTGGATCACTGTGTAGAGC    | 1139 |
| D  | b | 843  | GTCGGTCCCGAGGCCCTTGCTCAGCAACAACCTTTAGCTTGGGTCCACCATGCAGAAC    | 901  |
| <br>   |   |      |   |      |
| RESULT 13  |   |      |   |      |
| AR210322   |   |      |   |      |
| LOCUS  |   |      |   |      |
| DEFINITION Sequence 13 from patent US 6387657.   |   |      |   |      |
| ACCESSION AR210322   |   |      |   |      |
| VERSION AR210322.1 GI:21512523   |   |      |   |      |
| KEYWORDS   |   |      |   |      |
| SOURCE Unknown.  |   |      |   |      |
| ORGANISM Unclassified.   |   |      |   |      |
| REFERENCE 1 (bases 1 to 1293)  |   |      |   |      |
| AUTHORS Botstein,D.A., Cohen,R.L., Goddard,A.D., Gurney,A.L., Hillan,K.J., Lawrence,D.A., Levine,A.J., Pennica,D., Roy,MAnn. and Wood,W.I. |   |      |   |      |
| TITLE WSP polypeptides and nucleic acids encoding same   |   |      |   |      |
| JOURNAL Patent: US 6387657-A 13 14-MAY-2002;   |   |      |   |      |
| FEATURES   |   |      |   |      |
| source Location/Qualifiers   |   |      |   |      |
| BASE COUNT 232 a 425 c 393 g 243 t   |   |      |   |      |
| ORIGIN   |   |      |   |      |
| Query Match 32.9%; Score 561.4; DB 6; Length 1293;   |   |      |   |      |
| Best Local Similarity 78.0%; Pred. No. 2.2e-141;   |   |      |   |      |
| Matches 701; Conservative 0; Mismatches 196; Indels 2; Gaps 2;   |   |      |   |      |
| Q  | y | 243  | GGTGACATGAGGGCAGCCACTGATCCATCTTTGTGGCCACTTCTCTCTGCTTCTGCTTCTC | 302  |
| D  | b | 16   | GGGACATGAGAGGCACACCGAAGACCCACTCTTGCCCTTCTCCCTCTGCTGCTCCTC     | 75   |
| Q  | y | 303  | TCAATGGTGTGCCAGCTGTGCGGACACCCCTGTACCTGTCTTGGACACACACCCACG     | 362  |
| D  | b | 76   | TCAAAGGTGGCTACCCAGCTGTGCCGACACCATGTACCTGCCCTGGCCACCTCCCGA     | 135  |
| Q  | y | 363  | TGCCCCAGAGGGGTACCCCTGGTGGCTGTGGCTGTGGCTGTAAAGTGTGTGCACGG      | 422  |
| D  | b | 136  | TGCCCGTGGGAGTACCCCTGGTGGTGTGGCTGTGGCTGTGGCGGGTATGTGCACGG      | 195  |
| Q  | y | 423  | AGGCTGGGGAGTCTCGCACCACTGCATGTCTGGGACCCCGACCGAGGCGCTTGGTTGT    | 482  |
| D  | b | 196  | CGGCTGGGGAGCCCTCGCACCAACTCCACGTCTGCGACCGACCGAGGCGCTTGTCTG     | 255  |
| Q  | y | 483  | CAGCTGGGGCAGGCCCTGCGGCCATGCGGGCTGTGTCTCTTGGATGAGGATGACGGT     | 542  |
| D  | b | 256  | CAGCCCGGGCAGGACCCCGTGGCGGGGGCCCTGTGCCCTTTGGCAGAGGACGACGC      | 315  |
| Q  | y | 543  | AGCTGTGAGTGAATGGCCGACGTACCTGGATGGAGAGACCTTTAAACCCAATTGCAGG    | 602  |
| D  | b | 316  | AGCTGTGAGTGAACGGCCCGCTGTATCGGGAAGGGAGACCTTCCAGCCCCACTTGCAG    | 375  |
| Q  | y | 603  | TCTCTGTGCCCTGTGATGACGGTGTGCCTTCACTTGCTGCGCTGTGSCAGTAGGATGTG   | 662  |
| D  | b | 376  | ATCGCTGCCGCTGCGAGGACGGCGCTTCACTTGCTGCGCTGTGTCACGAGGATGTG      | 435  |
| Q  | y | 663  | CGGCTGCCAGCTGGGACTGCCACGCCCCAGAGAAATACAGTGGCCAGAAAGTGTCTGC    | 722  |
| D  | b | 436  | CGGCTGCCAGCTGGGACTGCCCCACCCACAGGAGGTCGAGGTCTCTGGGCAAGTGTCTG   | 495  |
| Q  | y | 723  | CCGAGTGGGTATGTGACCAGGGAGTGCACACCGCGATCCAGCGCTCCACGGCGCAGGA    | 782  |
| D  | b | 496  | CTGTAGTGGGTGTGCGGCCAAGGAGGGGAGTGGGACCCAGGCCCTTCCAGGCCCAAG     | 555  |
| Q  | y | 783  | CACCAACTTTCTGCCCTTGTCACTTCCTTGCTGTCTGTCTGTCTTGTCCAAATTTGAGC   | 842  |
| D  | b | 556  | CCCAAGTTTCTGGCCTTGTCTCTTCCCTGCCCTTGGTGTCCCTGTGCCCAGATGGAGC    | 615  |
| Q  | y | 843  | ACAGCCTGGGGCCCTCTCAACACACTGTGTGGCTGGGCATAGCCACCCCGAGTGTCCAAC  | 902  |









GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 23, 2003, 09:03:08 ; Search time 452.287 Seconds  
(without alignments)  
8504.365 Million cell updates/sec

Title: US-10-010-408-1  
Perfect score: 1708  
Sequence: 1 GAGCGTCTGATCTCCAGAG.....GCCTAGATAAACACCCAAA 1708

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_101002.\*

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Query | Score | Length | ID | Description |
|------------|-------|-------|--------|----|-------------|
| 1          | 1708  | 100.0 | 1708   | 20 | AAZ07516    |
| 2          | 1278  | 74.8  | 1734   | 20 | AAZ07516    |
| 3          | 753   | 44.1  | 753    | 20 | AAZ07517    |
| 4          | 681   | 39.9  | 681    | 20 | AAZ07521    |
| 5          | 657   | 38.5  | 753    | 20 | AAZ07516    |
| 6          | 583.4 | 34.2  | 1522   | 20 | AAZ07516    |
| 7          | 566.8 | 33.2  | 1337   | 22 | AAZ07516    |
| 8          | 566.6 | 33.2  | 1352   | 22 | AAZ07516    |
| 9          | 561.4 | 32.9  | 1266   | 21 | AAZ07516    |

|    |       |      |       |    |          |                    |
|----|-------|------|-------|----|----------|--------------------|
| 10 | 561.4 | 32.9 | 1266  | 22 | AAZ07516 | Human cDNA sequenc |
| 11 | 561.4 | 32.9 | 1266  | 22 | AAZ07516 | PRO261 coding sequ |
| 12 | 561.4 | 32.9 | 1266  | 22 | AAZ07516 | Human angiogenesis |
| 13 | 561.4 | 32.9 | 1285  | 19 | AAZ07516 | Human connective t |
| 14 | 561.4 | 32.9 | 1293  | 20 | AAZ07516 | Human WISP-2 prote |
| 15 | 561.4 | 32.9 | 1309  | 22 | AAZ07516 | Connective tissue  |
| 16 | 557   | 32.6 | 1267  | 21 | AAZ07516 | EGF-like homologue |
| 17 | 550.4 | 32.2 | 1267  | 21 | AAZ07516 | Human PRO261 cDNA  |
| 18 | 528.8 | 31.0 | 841   | 20 | AAZ07516 | Human WISP-2 prote |
| 19 | 508.4 | 29.8 | 750   | 20 | AAZ07516 | Human WISP-2 prote |
| 20 | 501.2 | 29.3 | 738   | 20 | AAZ07516 | Human WISP-2 prote |
| 21 | 408.2 | 23.9 | 439   | 24 | AAZ07516 | Rat OST23 gene fra |
| 22 | 210   | 12.3 | 210   | 20 | AAZ07516 | Rat HICP IGFBP dom |
| 23 | 203.4 | 11.9 | 2136  | 22 | AAZ07516 | Human full-length  |
| 24 | 195.8 | 11.5 | 13255 | 22 | AAZ07516 | Human immune/haema |
| 25 | 177   | 10.4 | 177   | 20 | AAZ07516 | Rat HICP VWC domai |
| 26 | 174   | 10.2 | 174   | 20 | AAZ07516 | Rat HICP TSP1 doma |
| 27 | 169.8 | 9.9  | 586   | 22 | AAZ07516 | Human foetal liver |
| 28 | 169.8 | 9.9  | 586   | 22 | AAZ07516 | Probe #6698 for ge |
| 29 | 169.8 | 9.9  | 586   | 22 | AAZ07516 | Human brain expres |
| 30 | 169.8 | 9.9  | 586   | 22 | AAZ07516 | Human bone marrow  |
| 31 | 169.8 | 9.9  | 586   | 22 | AAZ07516 | Probe #6485 for ge |
| 32 | 169.8 | 9.9  | 586   | 22 | AAZ07516 | Probe #8254 used t |
| 33 | 169.8 | 9.9  | 586   | 24 | AAZ07516 | Human genome-deriv |
| 34 | 163.8 | 9.6  | 2075  | 16 | AAZ07516 | Connective tissue  |
| 35 | 163.8 | 9.6  | 2075  | 18 | AAZ07516 | Human connective t |
| 36 | 163.8 | 9.6  | 2075  | 18 | AAZ07516 | Connective tissue  |
| 37 | 163.8 | 9.6  | 2075  | 19 | AAZ07516 | Human connective t |
| 38 | 163.8 | 9.6  | 2075  | 20 | AAZ07516 | Human connective t |
| 39 | 163.8 | 9.6  | 2075  | 21 | AAZ07516 | Human connective t |
| 40 | 163.8 | 9.6  | 2075  | 21 | AAZ07516 | Human connective t |
| 41 | 163.8 | 9.6  | 2075  | 22 | AAZ07516 | Human connective t |
| 42 | 163.8 | 9.6  | 2075  | 22 | AAZ07516 | Human connective t |
| 43 | 163.8 | 9.6  | 2075  | 22 | AAZ07516 | Human connective t |
| 44 | 163.8 | 9.6  | 2075  | 24 | AAZ07516 | Human benign prost |
| 45 | 163.8 | 9.6  | 2312  | 22 | AAZ07516 | Nucleotide sequenc |

#### ALIGNMENTS

RESULT 1  
AAZ07516

ID AAZ07516 standard; cDNA; 1708 BP.

XX AAZ07516;

XX AAZ07516;

XX 26-NOV-1999 (first entry)

XX Rat HICP polypeptide encoding cDNA.

XX Heparin-induced CCN-like protein; HICP; cell-associated activity; ss;

XX cardiovascular disorder; aberrant cell proliferation; fibrotic disorder.

XX Rattus sp.

XX WO9947556-A2.

XX 23-SEP-1999.

XX 18-MAR-1999; 99WO-US05999.

XX 19-MAR-1998; 98US-0044273.

XX (TUFT ) TUFTS COLLEGE.

XX Castellot JJ;

XX WPI; 1999-562060/47.

XX P-PSDB; AAZ27434.

XX Nucleic acid sequences encoding rat heparin-induced CCN-like protein, used in methods to identify modulators or in diagnostic applications



WNT-1 induced secreted protein; WISP-1, WISP-2, WISP-3, CTGF; tumour; connective tissue growth factor; cancer; melanoma; arteriosclerosis; leukaemia; lymphoid malignancy; haematopoiesis-related disorder; tissue-growth disorder; skin disorder; osteoporosis; fibrotic lesion; kidney disorder; bone-related disorder; osteoporosis; trauma; burn; connective tissue disorder; catabolic state; inflammation; testicular-related disorder; angiogenesis; immunological disorder; ss

Mus sp.  
WO9921998-A1.

06-MAY-1999.

29-OCT-1998: 98WO-US22991.

14-APR-1998; 98US-0081695.

29-OCT-1997; 97US-0063704.  
03-FEB-1998; 98US-0073612

1. NAME

(GETH) GENENTECH INC.

Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WI; Bolstein DA, Cohen RL, Goudard A, Gurney AL, Hillan K;

WPI; 1999-337420/28.

P-PSDB; AA11/651.

## New isolated Wnt-1 induced secreted poly

The present invention describes Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2 and WISP-3 have homology to connective tissue growth factor (CTGF). Products from the present invention can be used to treat WISP-related disorders such as breast, ovarian, and colon cancer or melanoma. The products can be used to treat arteriosclerosis. The products can also be used to treat other diseases e.g. benign and malignant tumours, leukaemia and lymphoid malignancies, neuronal, glial, astrocytal, hypothalamic and other glandular, macrophagal, epithelial, stromal, and blastocoeil disorders, haematopiasia-related disorders, tissue-growth disorders, skin disorders, desmoplasia, fibrotic lesions, kidney disorders, bone-related disorders such as osteoporosis, trauma such as burns, incisions, and other wounds, connective tissue disorders, catabolic states, testicular-related disorders, and inflammatory, angiogenic and immunologic disorders including arteriosclerosis. The products can also be used for detection and diagnosis especially of individuals with neoplastic cell growth or proliferation. The products can be used in the production of transgenic or knock-out animals. Antibodies can be used to induce death in Wisp-1, 2 or 3 overexpressing cells.

Sequence 1734 BP; 355 A; 491 C; 495 G; 393 T; 0 other:

Query Match. 74.8%; Score 1278; DB 20; Length 1734;  
Best Local Similarity 88.5%; Pred. No. 0;  
Matches 1520; Conservative 0; Mismatches 165; Indels 33;  
Gaps 11;

|    |     |  |     |
|----|-----|--|-----|
| Qy | 3   | CGCTCTGATCTCCAGAGGACCTGGGGTGGGACAGGGGCTTGGCAAGGCTGCAGCGC     | 62  |
| Db | 13  | CGCTCTGATCTCCAGAGGACCCGGGCTGGGACAGGGGCTTGGCGAGGCTGCAGTCG     | 72  |
| Qy | 63  | TG-GGCAGTGGCTTTGGAATGGAGTCTTTATTACTGGGACTGAGGAGCTAAGAGGCTCC  | 121 |
| Db | 73  | TGTGGCAGTAGCTTGGGATGGAGTCTTTCTGCTGGAACTGAGGAGCTGAGAGGCTCC    | 132 |
| Qy | 122 | TGTCAG----CTTGTCCTAAAGTCTTAGCACTTGTGTGGCTTGGGCTTCACACACTGTCA | 178 |
| Db | 133 | TGTCAGGCTCCTGTCTCTAACTTGGCACTTGGGTGGCTTGGGCTTCACACACTGTCA    | 192 |
| Qy | 179 | GACACCTTCGTGGGTCCACGGGCTCACCTTCAGTTTGAAGCTGGCTCCCAAGG        | 238 |

QY 1313 GTGCTGGGATAGTCAATGTTAGG-ACCAGACAGCAGATTCCTGAACTTCCAATTC 1371  
 DB 1327 ATGCTGGGATAGTGACAGTAATAGTACCAGGACGAGATTGCGCTGAAACATCCAGGTC 1386  
 QY 1372 CTTCTTGGACTTCGTATGCTTGTCCCAAGATGATGAATGAACTCGTAAGTGTACTT 1431  
 DB 1387 CTTCTTGGACTTCGTATGCTTGT-CCCAAGATGATGCTGACCTTGTAAAGTGTGCT 1445  
 QY 1432 TCCCTGACCTGAGACACCTCCCTGCTCGGGAAGTATTTCAGGGGACAGAAATCTCTGTGA 1491  
 DB 1446 TTCCTGATCTGAGACACCTCCCTCGGCTGGGA-----AGAAATTTCTGGGA 1492  
 QY 1492 ACATGAAGAGAT-GAATCACACTGCTCTTAAGAAATTCCTGAAAGTCCAGGAACCTTGAGC 1550  
 DB 1493 ACATGAAGAGATGGAATCACACTATTCTTAAGAGCGTTTGCCCAAGTCCAGGAACCTTGACC 1552  
 QY 1551 TTTGTATTTTTCAGGAATGCACATCTCTTAAGCACTCGGCAAAACAGAGGCTCCACACTT 1610  
 DB 1553 TTTGTATTTTGAATAATACATCTCTTAAGTCTCACAAGCAAG-AGGCTCCACACTT 1611  
 QY 1611 CTGGCAGGCGAGGCGCTTCTCTTCAGCATGAGAAAGCAAGGACAGAGTACTCTC 1670  
 DB 1612 CTGGCAGGCGAGGCGCTTCTCTTCAGCATGAGAGACAGAGTACTCTC 1671  
 QY 1671 CTCTGAGGACTAGTCTAGCCTAGAAATAAACACCCAAA 1708  
 DB 1672 CTCTGAGGACTGCGCGGCTCTGGAATAAACACCCAAA 1709

RESULT 3

AAZ07517  
 ID AAZ07517 standard; cdNA; 753 BP.  
 AC AAZ07517;  
 XX AAZ07517;  
 DT 26-NOV-1999 (first entry)  
 DE Rat HICP polypeptide coding sequence.  
 KW Heparin-induced CCN-like protein; HICP; cell-associated activity; ss;  
 KW cardiovascular disorder; aberrant cell proliferation; fibrotic disorder.  
 OS Rattus sp.  
 XX W09947556-A2.  
 XX 23-SEP-1999.  
 PF 18-MAR-1999; 99WO-US05999.  
 XX 19-MAR-1998; 98US-0044273.  
 PR (TUFT ) TUFTS COLLEGE.  
 PA Castellot JJ;  
 PI WPI; 1999-562060/47.  
 DR P-PSDB; AAV27434.  
 XX Nucleic acid sequences encoding rat heparin-induced CCN-like protein,  
 PT used in methods to identify modulators or in diagnostic applications -  
 PS Claim 5; Fig 1; 108pp; English.

CC The invention provides a rat heparin-induced CCN-like protein (HICP)  
 CC protein. Agents that stimulate or inhibit HICP protein activity or  
 CC expression, antisense HICP nucleic acid molecules and HICP antibodies,  
 CC can be used to modulate cell-associated activity. HICP protein activity  
 CC can be used to treat disorders characterized by aberrant HICP protein activity  
 CC or expression. Probes capable of hybridizing to HICP mRNA or antibodies  
 CC specific for HICP can be used to detect HICP activity in a biological  
 CC sample. HICP can be used to treat disorders, such as a cardiovascular or

CC fibrotic disorder, characterized by aberrant cell proliferation. The  
 CC present sequence represents the coding sequence of rat HICP.  
 XX Sequence 753 BP; 132 A; 235 C; 230 G; 156 T; 0 other;  
 SQ Query Match 44.1%; Score 753; DB 20; Length 753;  
 Best Local Similarity 100.0%; Pred. No. 3.8e-216;  
 Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 249 ATGAGGGGAGCCCACTGATCCATCTCTGGCCACTTCTCTCTCTCTCTCTCTCAATG 308  
 DB 1 ATGAGGGGAGCCCACTGATCCATCTCTGGCCACTTCTCTCTCTCTCTCTCTCTCAATG 60  
 QY 309 GTGTGTGCCCACTGTGCCGACACCTGTACTCTGTCTTGGACACCCAGTGCCTCAATG 368  
 DB 61 GTGTGTGCCCACTGTGCCGACACCTGTACTCTGTCTTGGACACCCAGTGCCTCAATG 120  
 QY 369 CAGGGGGTACCCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 428  
 DB 121 CAGGGGGTACCCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 180  
 QY 429 GGGGAGTCTGGGACACCTGTCTGCGACCCAGCAGGGCTGGTGGTGGTGGTGGTGGTGG 488  
 DB 181 GGGGAGTCTGGGACACCTGTCTGCGACCCAGCAGGGCTGGTGGTGGTGGTGGTGGTGG 240  
 QY 489 GGGGAGTCTGGGACACCTGTCTGCGACCCAGCAGGGCTGGTGGTGGTGGTGGTGGTGG 548  
 DB 241 GGGGAGTCTGGGACACCTGTCTGCGACCCAGCAGGGCTGGTGGTGGTGGTGGTGGTGG 300  
 QY 549 GAGGTGAATGGCGGAGGTACCTGGATGGAGAGACCTTTAAACCCAAATTCAGGGTCTCG 608  
 DB 301 GAGGTGAATGGCGGAGGTACCTGGATGGAGAGACCTTTAAACCCAAATTCAGGGTCTCG 360  
 QY 609 TGCCGCTGTGATGACGGTGGCTTACCTGCTGCGCTGCTGAGTGGAGTGGAGTGGAGTGG 668  
 DB 361 TGCCGCTGTGATGACGGTGGCTTACCTGCTGCGCTGCTGAGTGGAGTGGAGTGGAGTGG 420  
 QY 669 CCCAGCTGGGAGTGGCGGACGCCCCAAGAGAAATACAGTGGCCAGGAAAGTGGTGGCCCGAG 728  
 DB 421 CCCAGCTGGGAGTGGCGGACGCCCCAAGAGAAATACAGTGGCCAGGAAAGTGGTGGCCCGAG 480  
 QY 729 TGGGTATGACCGGAGGTGACACCGGGGATCCAGGGTCCAGCGGCGGAGGACACCA 788  
 DB 481 TGGGTATGACCGGAGGTGACACCGGGGATCCAGGGTCCAGCGGCGGAGGACACCA 540  
 QY 789 CTTTCTGCTTGTCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 848  
 DB 541 CTTTCTGCTTGTCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
 QY 849 TGGGGCCCTGCTCAACCACTGTGGGCTGGGCTAGCCACCCGAGTGTCCAAACAGAAC 908  
 DB 601 TGGGGCCCTGCTCAACCACTGTGGGCTGGGCTAGCCACCCGAGTGTCCAAACAGAAC 660  
 QY 909 CGATTCTGCCAATCTGAGATCCAAACCGCCCTGTGTCTGCGCCAGACCCCTGCTGCGCAGCC 968  
 DB 661 CGATTCTGCCAATCTGAGATCCAAACCGCCCTGTGTCTGCGCCAGACCCCTGCTGCGCAGCC 720  
 QY 969 AGGAGCCACAGCTCATGGAACAGTGTCTTCTAA 1001  
 DB 721 AGGAGCCACAGCTCATGGAACAGTGTCTTCTAA 753

RESULT 4

AAZ07521  
 ID AAZ07521 standard; cdNA; 681 BP.  
 XX AAZ07521;  
 AC AAZ07521;  
 DT 26-NOV-1999 (first entry)  
 DE Rat HICP mature polypeptide coding sequence.  
 XX Heparin-induced CCN-like protein; HICP; cell-associated activity; ss;

KW cardiovascular disorder; aberrant cell proliferation; fibrotic disorder.  
 OS Rattus sp.  
 XX WO9947556-A2.  
 XX 23-SEP-1999.  
 XX 18-MAR-1999; 99WO-US05999.  
 XX 19-MAR-1998; 98US-0044273.  
 XX (TUFT ) TUFTS COLLEGE.  
 XX Castellot JJ;  
 XX WPI; 1999-562060/47.  
 XX P-PSDB; AAY27440.  
 XX Nucleic acid sequences encoding rat heparin-induced CCN-like protein,  
 PT used in methods to identify modulators or in diagnostic applications -  
 XX Disclosure; Fig 2; 108pp; English.  
 XX The invention provides a rat heparin-induced CCN-like protein (HICP)  
 CC protein. Agents that stimulate or inhibit HICP protein activity or  
 CC expression, antisense HICP nucleic acid molecules and HICP antibodies,  
 CC can be used to modulate cell-associated activity. HICP modulators can be  
 CC used to treat disorders characterized by aberrant HICP protein activity  
 CC or expression. Probes capable of hybridizing to HICP mRNA or antibodies  
 CC specific for HICP can be used to detect HICP activity in a biological  
 CC sample. HICP can be used to treat disorders, such as a cardiovascular or  
 CC fibrotic disorder, characterized by aberrant cell proliferation. The  
 CC present sequence represents the coding sequence of rat HICP mature  
 CC polypeptide.  
 XX  
 XX Sequence 681 BP; 121 A; 211 C; 215 G; 134 T; 0 other;  
 SQ

Query Match 39.9%; Score 681; DB 20; Length 681;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-194;  
 Matches 681; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 318 CAGCTGTGCGGACACCTGTACTGTCTTGGACACACCCAGTGCACAGGGGGTA 377  
 DB 1 CAGCTGTGCGGACACCTGTACTGTCTTGGACACACCCAGTGCACAGGGGGTA 60  
 QY 378 CCCCTGTGTGATGCTGTGCTGTGTAAGTGTGTACGAGGCTGGGGAGTCC 437  
 DB 61 CCCCTGTGTGATGCTGTGCTGTGTAAGTGTGTACGAGGCTGGGGAGTCC 120  
 QY 438 TGGGACACCTGTCTGCGACCCAGCGGCTGTGTTGTGAGCTGGGGCAGGC 497  
 DB 121 TGGGACACCTGTCTGCGACCCAGCGGCTGTGTTGTGAGCTGGGGCAGGC 180  
 QY 498 CTTGGGGCCATGGGGTGTGTCTTGGATGAGATGACGGTACGTGTGAGTGAAT 557  
 DB 181 CTTGGGGCCATGGGGTGTGTCTTGGATGAGATGACGGTACGTGTGAGTGAAT 240  
 QY 558 GCGCGGAGTACCTGGATGAGAGACCTTTAAACCAATTCAGGCTGCTCCCGCTGT 617  
 DB 241 GCGCGGAGTACCTGGATGAGAGACCTTTAAACCAATTCAGGCTGCTCCCGCTGT 300  
 QY 618 GATGAGGTTGGCTTCACTGCGGCTGTGTCAGTGAAGATGTGGGGTGGCCAGCTGG 677  
 DB 301 GATGAGGTTGGCTTCACTGCGGCTGTGTCAGTGAAGATGTGGGGTGGCCAGCTGG 360  
 QY 678 GATGAGGTTGGCTTCACTGCGGCTGTGTCAGTGAAGATGTGGGGTGGCCAGCTGG 737  
 DB 361 GATGAGGTTGGCTTCACTGCGGCTGTGTCAGTGAAGATGTGGGGTGGCCAGCTGG 420  
 QY 738 GACGAGGAGTACACGGGATGTCAGGCTTCCAGCGGCAAGGACACCAACTTTCGCC 797  
 DB 421 GACGAGGAGTACACGGGATGTCAGGCTTCCAGCGGCAAGGACACCAACTTTCGCC 480

QY 798 CTTGTCACTCCTCCTCTGCTGATGCTCTTCCAAATTTGGAGCAGCCTGGGGCCCC 857  
 DB 481 CTTGTCACTCCTCCTCTGCTGATGCTCTTCCAAATTTGGAGCAGCCTGGGGCCCC 540  
 QY 858 TGTCTAACCACTGTGGGTGGGCATAGCCACCCGAGTGTCCAAACAGAACCCGATTCTGC 917  
 DB 541 TGTCTAACCACTGTGGGTGGGCATAGCCACCCGAGTGTCCAAACAGAACCCGATTCTGC 600  
 QY 918 CAACTGGAGATCCAAACGCGCCTGTGCTGCCAGACCCCTGCTGGCAGCCAGGAGCCAC 977  
 DB 601 CAACTGGAGATCCAAACGCGCCTGTGCTGCCAGACCCCTGCTGGCAGCCAGGAGCCAC 660  
 QY 978 AGCTCATGGAACAGTGTTC 998  
 DB 661 AGCTCATGGAACAGTGTTC 681

RESULT 5  
 AAX76489/c  
 ID AAX76489 standard; DNA: 753 BP.  
 XX  
 AC AAX76489;  
 XX  
 DT 06-AUG-1999 (first entry)  
 XX Mouse WISP-2 protein complementary nucleotide sequence SEQ ID NO:18.  
 DE  
 XX WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;  
 KW connective tissue growth factor; cancer; melanoma; arteriosclerosis;  
 KW leukaemia; lymphoid malignancy; haematopoiesis-related disorder;  
 KW tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;  
 KW kidney disorder; bone-related disorder; osteoporosis; trauma; burn;  
 KW connective tissue disorder; catabolic state; inflammation;  
 KW testicular-related disorder; angiogenesis; immunological disorder; SS.  
 XX  
 OS Mus sp.  
 XX  
 XX WO9921998-AL.  
 XX  
 XX 06-MAY-1999.  
 XX  
 XX 29-OCT-1998; 98WO-US22991.  
 XX  
 XX 14-APR-1998; 98US-0081695.  
 PR 29-OCT-1997; 97US-0063704.  
 PR 03-FEB-1998; 98US-0073612.  
 XX  
 XX (GETH ) GENENTECH INC.  
 XX  
 XX Botstein DA, Cohen RL, Goddard A, Gurney AL, Hillan K;  
 PI Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WI;  
 XX  
 XX WPI; 1999-337420/28.  
 DR  
 XX New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3  
 PT Disclosure; Page 179-180; 284pp; English.  
 XX  
 XX The present invention describes Wnt-1 induced secreted polypeptides,  
 CC WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2  
 CC and WISP-3 have homology to connective tissue growth factor (CTGF).  
 CC Products from the present invention can be used to treat WISP-related  
 CC disorders such as breast, ovarian, and colon cancer or melanoma. The  
 CC products can be used to treat arteriosclerosis. The products can also be  
 CC used to treat other diseases e.g. benign and malignant tumours,  
 CC leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,  
 CC hypothalamic and other glandular, macrophagal, epithelial, stromal, and  
 CC blastocoeleic disorders, haematopoiesis-related disorders, tissue-growth  
 CC disorders, skin disorders, desmoplasia, fibrotic lesions, kidney  
 CC disorders, bone-related disorders such as osteoporosis, trauma such as  
 CC burns, incisions, and other wounds, connective tissue disorders,  
 CC catabolic states, testicular-related disorders, and inflammatory,

CC angiogenic and immunologic disorders including arteriosclerosis. The  
CC products can also be used for detection and diagnosis especially of  
CC individuals with neoplastic cell growth or proliferation. The products  
CC can be used in the production of transgenic or knock-out animals.  
CC Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing  
CC cells.

XX Sequence 753 BP; 159 A; 229 C; 238 G; 127 T; 0 other;

```
Query Match      38.5%; Score 657; DB 20; Length 753;
Best Local Similarity 93.0%; Pred. No. 3.2e-187;
Matches 700; Conservative 0; Mismatches 50; Indels 3; Gaps 1;

QY 249 ATGAGGGGACCCCACTGATCCATCTTCTGGCCCACTTCTTCCCTGCTGCTTCTCTCAATG 308
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 753 ATGAGGGGCAACCCCACTGATCCATCTTCTGGCCCACTTCTTCTCTGCAATCTCTCAATG 694
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 309 GTGTGTGCCAGCTGTGCGGACACCCCTGTACCTGTCTTGGACACACCCCACTGCCCCA 368
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 693 GTGTATTTCCCACTGTGCGGACACCCCTGTGCTGTCTTGGACACACCCCACTGCCCCA 634
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 369 CAGGGGTACCCCTGGTCTGTGATGGCTGTGCTGCTGTAAGTGTGTGCAGGAGGCTG 428
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 633 CCGGGGTACCCCTGGTCTGTGATGGCTGTGCTGCTGTGCTGTGCTGCTGCTGCTGCTG 574
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 429 GGGGAGTCTCGACCACTGCATGTCTGCGACCCCACTGAGGGCTGTGTTTGTGAGCCT 488
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 573 GGGGAGTCTCGACCACTGCATGTCTGCGACCCCACTGAGGGCTGTGTTTGTGAGCCT 514
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 489 GGGGAGGCTCTGGGGCCATGGGCTGTGTCTCTTGGATGAGGATGACGTAGCTGT 548
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 513 GGGGAGGCTCTGGGGCCATGGGCTGTGTCTCTTGGATGAGGATGACGTAGCTGT 454
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 549 GAGGTGAATGCGCGAGTACTGTGATGGAGAGACCTTTAAACCAATTCAGGCTCTG 608
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 453 GAGGTGAATGCGCGAGTACTGTGATGGAGAGACCTTTAAACCAATTCAGGCTCTG 394
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 609 TGCCGCTGTATGAGCGGTGGCTTACCTGCGCTGTGCGAGTGTGAGGATGTGCGGCTG 668
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 393 TGCCGCTGTATGAGCGGTGGCTTACCTGCGCTGTGCGAGTGTGAGGATGTGCGGCTG 334
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 669 CCCAGCTGGGACTGCCAGCCCCCAAGAGATACAGGTGCGAGGAAAGTGTGCCCCAG 728
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 333 CCCAGCTGGGACTGCCAGCCCCCAAGAGATACAGGTGCGAGGAAAGTGTGCCCCAG 274
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 729 TGGGTATGTACCGGGAGTGA--CACCAGCGATCCAGCGCTCCACCGCGCAAGGACAC 785
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 273 TGGGTGTGTACCGAGCATGTATGACGCGGCAATCCAGCCCTCTCAGCCCCAAGGACAC 214
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 786 CAACCTTCTGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 845
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 213 CAACCTTCTGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 154
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 846 GCCTGGGCGCCCTGCTCAACCACTGTGGCTGGGCTAGCCAGCCCGAGTGTCCAAACAG 905
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 153 GCCTGGGCGCCCTGCTCAACCACTGTGGCTGGGCTAGCCAGCCCGAGTGTCCAAACAG 94
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 906 AACCGATTCTGCCAACTGGAGATCCAAACCGCGCTGTGCTGCGCAGACCCCTGCTGGCA 965
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 93 AACCGATTCTGCCAACTGGAGATCCAGGCTGCGCTGTGCTGCTGCTGCTGCTGCTGCTG 34
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 966 GCCAGGAGCCACAGCTCATGGAACAGTGTCTTTC 998
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 33 TCCAGGAGCCAGGCTCATGGAACAGTGTCTTTC 1
```

RESULT 6

AAAX16595

ID AAAX16595 standard; cDNA; 1522 BP.

XX

AC AAAX16595;

XX

DT 29-APR-1999 (first entry)

XX Human growth factor-like protein encoding cDNA.  
DE Human: growth factor-like protein; HGFLP; flsp-12; CTGF; PDGF family;  
KW connective tissue growth factor; platelet-derived growth factor; keloid;  
KW connective tissue disorder; cancer; ankylosing spondylitis; scleroderma;  
KW atherosclerosis; Dupuytren's contracture; eosinophilic fasciitis;  
KW Felt syndrome; Goodpasture's disease; Hunter syndrome; Hurler syndrome;  
KW Marfan syndrome; nodular fasciitis; osteogenesis imperfecta; restenosis;  
KW rheumatoid arthritis; systemic lupus erythematosus; ss.  
XX Homo sapiens.  
OS  
XX  
XX Key Location/Qualifiers  
FH 256..1008  
FT /\*tag= a  
XX  
XX WO9858063-A1.  
XX 23-DEC-1998.  
XX 18-JUN-1998; 98WO-US12787.  
XX 19-JUN-1997; 97US-0878990.  
XX (INCY-) INCYTE PHARM INC.  
XX Corley NC, Hillman JL, Lal P;  
XX  
XX WPI; 1999-080954/07.  
XX P-PSDB; AAW94616.  
XX New polynucleotide encoding growth factor-like protein - useful for  
XX preventing and treating connective tissue disorders and cancer  
XX  
XX Claim 4; Fig 1; 63pp; English.  
XX  
XX The present sequence encodes human growth factor-like protein (HGFLP)  
XX which has homology to connective tissue growth factor (CTGF) and flsp-12,  
XX which are members of the platelet-derived growth factor (PDGF)  
XX superfamily of growth factors. Antisense oligonucleotides and HGFLP  
XX antagonists can be used to treat or prevent cancer, and connective  
XX tissue disorders including ankylosing spondylitis, atherosclerosis,  
XX Dupuytren's contracture, eosinophilic fasciitis, Felt syndrome,  
XX Goodpasture's disease, Hunter syndrome, Hurler syndrome, keloids, Marfan  
XX syndrome, nodular fasciitis, osteogenesis imperfecta, polyarthritits  
XX nodosa, rheumatoid arthritis, scleroderma, systemic lupus erythematosus,  
XX and restenosis following angioplasty. HGFLP antibodies and  
XX polynucleotides can also be used in diagnostic assays for conditions or  
XX diseases characterised by GRFLP expression.

XX Sequence 1522 BP; 279 A; 496 C; 459 G; 288 T; 0 other;

```
Query Match      34.2%; Score 583.4; DB 20; Length 1522;
Best Local Similarity 77.6%; Pred. No. 6.9e-165;
Matches 747; Conservative 0; Mismatches 206; Indels 10; Gaps 3;

QY 167 CACACACTGTGACACACCTTGTGTGGCTCCAGCGCTCACCTTCAGGTTTGAAGCTG 226
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 182 CACACAGGACAGGACACCCCTTGTGTGGCTTCACAGTTTCACCTTCAGGCTCAAGCTG 241
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 227 GCTCCACAAGGACACGCTGATGAGGGGAGCGCCACTGATCCATCTTCTGGCCACTTC 286
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 242 GCTCTGCAAGG-----GACATGAGAGGACACCCAGAACCCACCTCCTGGCCTTCTC 293
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 287 CTTCTCTGCTTCTCTCAATGGTGTGTGCCAGCTGTGCCGACACCCCTGTACCTGTCC 346
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 294 CTTCTCTGCTTCTCTCAAAAGGTGCGTACCCAGCTGTGCCGACACCATGTACCTGCC 353
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 347 TTGGACACACCCAGTGTCCACAGGGGTATCCCTGTGTGTGATGCTGTGGCTGTGCTG 406
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 354 CTGCCCACTCCCGATGCCCGCTGGGAGTACCCCTGTGTGATGCTGTGGCTGTGCTG 413
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

|   |      |     |   |      |
|---|------|-----|---|------|
| QY  | 407  | TA  | AAGTGTGTGCACGAGGAGCTGGGGAGTCTCGGACCACTGCATGTCTGCGACCCCGAG   | 466  |
|   |      |     |   |      |
| Db  | 414  | CC  | GGGTATGTGCAACGGCGCTGGGGAGCCCTCGGACCACTTCAGGTCTGCGACGCCAG    | 473  |
|   |      |     |   |      |
| QY  | 467  | CA  | AGGCGCTGGTTGTTCAGCTCGGGCAGGCGCTCGGCGCATGGGCGCTGTGTGTCTCTT   | 526  |
|   |      |     |   |      |
| Db  | 474  | CC  | AGGCGCTGGTCTGCCAGCCCGGGCAGACCCGGTGGCGGGGGCGCCCTGTGCGCTCTT   | 533  |
|   |      |     |   |      |
| QY  | 527  | GG  | ATGAGGATGACGGTAGCTGTGAGGTGAATGGCCGACAGTACTGCTGATGGAGAGACCTT | 586  |
|   |      |     |   |      |
| Db  | 534  | GG  | CAGAGGACGACAGCAGCTGTGAGGTGAACGCCCGCTGTATCGGAAGGGGAGACCTT    | 593  |
|   |      |     |   |      |
| QY  | 587  | TA  | AACCAATTGACAGGGTCTGTGCGCGCTGTGTATGACGGTGGCTTCACCTGCTCGCGCT  | 646  |
|   |      |     |   |      |
| Db  | 594  | CC  | AGCCCACTGCAAGCATCCGTGCGCGCTCGAGAGACGGCGGCTTCACCTGCGCGCT     | 653  |
|   |      |     |   |      |
| QY  | 647  | GT  | CACTGAGGATGTCGGGCTGCCAGCTGGGACTGCCACGCCCCCAAGAGAATACAGGT    | 706  |
|   |      |     |   |      |
| Db  | 654  | GT  | CACGAGGATGTGCGGCTGCCAGCTGGGACTGCCCCACCCAGGAGGTTCAGGT        | 713  |
|   |      |     |   |      |
| QY  | 707  | GC  | CAGGAAATGCTGCCCGAGTGGGTATGTGACCAAGGAGTGACACCGGCGATCCAGCG    | 766  |
|   |      |     |   |      |
| Db  | 714  | CT  | TGGCAAGTGTGCCCTGAGTGGGTGTGCGGCCAAGGAGGGGACTGGGACCCAGCC      | 773  |
|   |      |     |   |      |
| QY  | 767  | CT  | CCAGCGGCAAGGACACCAACTTTCTGCGCTTTGTCACTTCCTGCGCTGATGCTCC     | 826  |
|   |      |     |   |      |
| Db  | 774  | CT  | TCCAGCCCAAGGACCCAGTTTCTGGCCCTTGTCTCTCCCTGCCCCCTGGTGCTCC     | 833  |
|   |      |     |   |      |
| QY  | 827  | TT  | GTCACAAATTGAGACACAGCCCTGGGCGCCCTGCTCAACACACCTGTGGCTGGCATAGC | 886  |
|   |      |     |   |      |
| Db  | 834  | CT  | CCCAAGATGGAGACAGGCGCTGGGACCCCTGCTCGACCACTGTGGGTGGGCGATGGC   | 893  |
|   |      |     |   |      |
| QY  | 887  | CA  | CCCGAGTGTCCAACCAAGAACCGATTCTGCCAACTGGAGATCCAACGCCCGCTGTCT   | 946  |
|   |      |     |   |      |
| Db  | 894  | CA  | CCCGGTGTCCAACCAAGAACCGCTTCTGCCACTGGAGACCCAGCGCGCTGTGCT      | 953  |
|   |      |     |   |      |
| QY  | 947  | GCC | CAGACCTGCTGGCAGCCAGAGCCACAGCTCATGGAACAGTGTCTTCTA-AGGCC      | 1005 |
|   |      |     |   |      |
| Db  | 954  | GT  | CAGCGCCCTGCCACCTTCCAGGGGTGCGATCCACAAACAGTGCCTTCTAGAGCCG     | 1013 |
|   |      |     |   |      |
| QY  | 1006 | AA  | CTGGGATGCGGATACAGGGCTGCCATCTTCAGCAATGACCCCTAGACACAGGCCCT    | 1065 |
|   |      |     |   |      |
| Db  | 1014 | GG  | CTGGAAATGGGGACAGGTGTCCACCATCCCCAGCTGGTGGCCCTGTGCTGGGCCCT    | 1073 |
|   |      |     |   |      |
| QY  | 1066 | GG  | ACTGCTGTGTAGTGTCTCTTCTCCATGCTTTGGCTGCAGTTAACTGTCTCTTGA      | 1124 |
|   |      |     |   |      |
| Db  | 1074 | GG  | CTCATGGAAGATGGTCCGTGCCAGGCCCTTGGCTGCAGGCACACTTTAGCTTGGG     | 1133 |
|   |      |     |   |      |
| QY  | 1125 | TT  | C 1127  |      |
|   |      |     |   |      |
| Db  | 1134 | TT  | C 1136  |      |
|   |      |     |   |      |
| RESULT 7  |      |     |   |      |
| AAH46952  |      |     |   |      |
| ID ID AAH46952 standard; cdna; 1337 BP.                     |      |     |   |      |
| XX AAH46952;  |      |     |   |      |
| XX AC   |      |     |   |      |
| XX AC   |      |     |   |      |
| XX TT   |      |     |   |      |
| XX 25-SEP-2001 (first entry)                                |      |     |   |      |
| XX Human secreted protein encoding cdna (clone id HB0DB48). |      |     |   |      |
| XX DE   |      |     |   |      |
| XX KW   |      |     |   |      |
| XX KW   |      |     |   |      |
| XX KW   |      |     |   |      |
| XX KW   |      |     |   |      |
| XX KW   |      |     |   |      |
| XX OS   |      |     |   |      |
| XX Homo sapiens.  |      |     |   |      |
| XX PN   |      |     |   |      |
| XX WO200155430-A1.  |      |     |   |      |
| XX PD   |      |     |   |      |
| XX 02-AUG-2001.   |      |     |   |      |

[illegible]







OY 1082 CT-CTTCTCCAGCTCTTGGCTGCGATTAACTGCTCTGCTTGGATTCACTGCTGAGAGC 1139  
 Db 844 GTCCGTGCCAGGCCCTTGGCTGCGAGGCAACACTTTAGCTTGGGTCCACCATGCAGAAC 902

RESULT 10  
 AAS21403  
 ID AAS21403 standard; cDNA: 1266 BP.  
 XX AAS21403;  
 AC AAS21403;  
 DT 24-OCT-2001 (first entry)  
 XX  
 DE Human cDNA sequence encoding for PRO261 polypeptide.  
 KW Human secretory and transmembrane; PRO: mammalian; cancer; lung;  
 KW breast; prostate; cervical; tumour necrosis factor-alpha; TNF-alpha;  
 KW cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;  
 KW adipocyte; A-peptide; factor VIIA; gene therapy; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200140466-A2.  
 XX  
 PD 07-JUN-2001.  
 XX  
 PF 01-DEC-2000; 2000WO-US32678.  
 XX  
 PR 01-DEC-1999; 99WO-US28301.  
 PR 01-DEC-1999; 99WO-US28634.  
 PR 02-DEC-1999; 99WO-US28551.  
 PR 02-DEC-1999; 99WO-US28564.  
 PR 02-DEC-1999; 99WO-US28565.  
 PR 09-DEC-1999; 99US-0170262.  
 PR 16-DEC-1999; 99WO-US30095.  
 PR 20-DEC-1999; 99WO-US30911.  
 PR 20-DEC-1999; 99WO-US30999.  
 PR 30-DEC-1999; 99WO-US1243.  
 PR 06-JAN-2000; 2000WO-US00277.  
 PR 06-JAN-2000; 2000WO-US00376.  
 PR 11-FEB-2000; 2000WO-US03565.  
 PR 18-FEB-2000; 2000WO-US04341.  
 PR 18-FEB-2000; 2000WO-US04342.  
 PR 22-FEB-2000; 2000WO-US04414.  
 PR 24-FEB-2000; 2000WO-US04914.  
 PR 24-FEB-2000; 2000WO-US05004.  
 PR 01-MAR-2000; 2000WO-US05601.  
 PR 20-MAR-2000; 2000WO-US07377.  
 PR 21-MAR-2000; 2000WO-US07532.  
 PR 30-MAR-2000; 2000WO-US08439.  
 PR 17-MAY-2000; 2000WO-US13705.  
 PR 22-MAY-2000; 2000WO-US14042.  
 PR 30-MAY-2000; 2000WO-US14941.  
 PR 02-JUN-2000; 2000WO-US15264.  
 PR 10-NOV-2000; 2000WO-US30873.  
 XX  
 (GETH ) GENENTECH INC.  
 XX  
 PI Baker KP, Beresini M, DeForge L, Desnoyers L, Filvaroff E, Gao W;  
 PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;  
 PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WL, Zhang Z;  
 XX  
 DR WPI; 2001-408281/43.  
 DR P-PSDB; AAU12331.  
 XX  
 PT Isolated, secretory and transmembrane PRO polypeptide used to detect  
 PT other PRO polypeptides, link bioactive molecules to cells expressing  
 PT PRO polypeptides, and detect the presence of mammalian tumours e.g.  
 PT lung, breast, prostate, cervical  
 XX  
 PS Claim 3; Fig 319; 813pp; English.  
 XX  
 CC AAS21244-AAS21518 encode for novel human secretory and transmembrane

CC PRO polypeptides. The PRO polypeptides are useful to detect other  
 CC PRO polypeptides, to link bioactive molecules to cells expressing  
 CC PRO polypeptides, to modulate biological activities of cells expressing  
 CC PRO polypeptides, and to detect the presence of mammalian lung, colon,  
 CC breast, prostate, rectal, cervical or liver tumours by comparing PRO  
 CC polypeptide expression in a cell sample to that in a control sample.  
 CC Some of the 275 sequences are also useful to stimulate the release of  
 CC tumour necrosis factor-alpha (TNF-alpha) from human blood, the  
 CC proliferation or differentiation of chondrocytes, the proliferation or  
 CC gene expression in pericyte cells, the release of proteoglycans from  
 CC cartilage, the proliferation of inner ear utricular supporting cells or  
 CC of T-lymphocytes, the release of a cytokine from peripheral blood  
 CC monocytes (PMCs), or the proliferation of endothelial cells. Some of  
 CC the PRO polypeptides may modulate glucose or free fatty acid uptake by  
 CC skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide  
 CC to factor VIIA. The PRO polypeptides can be used in assays to identify  
 CC molecules involved in binding interactions. The polynucleotides encoding  
 CC PRO polypeptides can be used to generate probes, antisense RNA/DNA,  
 CC transgenic or knock out animals and can be used in gene therapy.  
 XX  
 SQ Sequence 1266 BP; 216 A; 418 C; 390 G; 242 T; 0 other;  
 Query Match 92.9%; Score 561.4; DB 22; Length 1266;  
 Best Local Similarity 78.0%; Pred. No. 2.7e-158;  
 Matches 701; Conservative 0; Mismatches 196; Indels 2; Gaps 2;  
 OY 243 GTGACATGAGGCGCAGCCACTGATCCATCTTCTGGCCACTTCTCTCTGCTTCTC 302  
 Db 4 GGGACATGAGAGGCGACACACCAAGACCCACTCTCTGGCTTCTCTCTCTCTCTC 63  
 OY 303 TCAATGTGTGTGCCAGCTGTGGGAGACACCTGTACTGTCTTCTGTGACACCCAG 362  
 Db 64 TCAAAGGTGCGTACCCAGCTGTGCCGACACCATGTACTGTCCCTGTGCCACCTCCGCA 123  
 OY 363 TGCCACAGGGGTACCCCTGGTGTGATGGCTGTGGCTGTCTGTAAGTGTGTGCAGG 422  
 Db 124 TGCCGCTGGGAGTACCCCTGGTGTGATGGCTGTGGCTGTCTGCGGGGTATGTGCAGG 183  
 OY 423 AGGCTGGGGAGTCTGCGACACCTGATGCTGCGACCCAGCCAGCGGCTGTTTGT 482  
 Db 184 CGGCTGGGGAGGCTGCGACCACTCCACGCTCTGCGACGCCAGCGGCTGTCTGC 243  
 OY 483 CAGCTGGGGAGGCTTGGCGGCATGGGCTGTGTCTCTCTTGGATGAGATGACGGT 542  
 Db 244 CAGCCCGGGCAGGACCCGGTGGCGGGGGCCCTGTGCTCTTGGCAGAGCAGCAGC 303  
 OY 543 AGCTGTAGGTCAATGGCGCCAGGTTACCTGATGGAGAGACTTTAAACCAATTGCAGG 602  
 Db 304 AGCTGTAGGTGAACGGCCGCTGTATCGGGAAGGGAGAGCTTCCAGCCCACTGCAGC 363  
 OY 603 GTCCTGTCCGCTGTGATGAGCGGTGGCTTACCTGCTGCGGCTGTGAGTGTGAGTGTG 662  
 Db 364 ATCCGCTGCCGCTGGAGGAGCGGGCTTACCTGCTGCGGCTGTGAGCAGGATGTG 423  
 OY 663 CGGCTGCCAGCTGGGACTGCCAGCCCCAGAGAAATACAGGTGCCAGGAAAGTGTGTC 722  
 Db 424 CGGCTGCCAGCTGGGACTGCCAGCCCCAGCAGGAGGTCGAGGTCTTGGGCAAGTGTGC 483  
 OY 723 CCCAGTGGGTATGTGACCAAGGAGTGACACCGGCGATCCAGCGCTCCACGCGCAAGA 782  
 Db 484 CTTGAGTGGGTGTGGGCAAGGAGGGGAGTGGGGACCCAGCCCTTCCAGCCCAAGA 543  
 OY 783 CACCAACTTCTGCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 842  
 Db 544 CCCAGTGGGTGTGGGCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 603  
 OY 843 ACAGCTGGGCGCGCTGCTCAACCACTGTGGGTGGGCATAGCAGCCAGTGTCTCAAC 902  
 Db 604 ACGGCTGGGAGCGGCTGCTCGACCACTGTGGGTGGGCATGGCCAGCCAGGCTGTCTCA 663  
 OY 903 CAGAACCATTTCTGGAACCTGGAGATCCAAAGCCGCTGTCTGCTGCCAGACCCCTGCTG 962  
 Db 664 CAGAACCCCTTCTGCGGAGTGGAGAGCCAGCGCCGCTGTGCTGCTGCTGCTGCTGCTG 723















GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 15:55:06 ; Search time 92.9244 Seconds  
(without alignments)  
5636.882 Million cell updates/sec

Title: US-10-010-408-1  
Perfect score: 1708  
Sequence: 1 GACGCTCTGATCTCCAGAG.....GCTAGATAAACACCCAAA 1708

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_NA:\*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description      |
|------------|-------|-------------|--------|----|------------------|
| 1          | 1278  | 74.8        | 1734   | 4  | US-09-182-145-17 |
| 2          | 1278  | 74.8        | 1734   | 4  | US-09-182-145-18 |
| 3          | 561.4 | 32.9        | 1293   | 4  | US-09-182-145-13 |
| 4          | 561.4 | 32.9        | 1293   | 4  | US-09-182-145-14 |
| 5          | 528.8 | 31.0        | 841    | 4  | US-09-182-145-39 |
| 6          | 501.2 | 29.3        | 738    | 4  | US-09-182-145-38 |
| 7          | 163.8 | 9.6         | 2075   | 1  | US-08-167-628-1  |
| 8          | 163.8 | 9.6         | 2075   | 1  | US-08-386-680-1  |
| 9          | 163.8 | 9.6         | 2075   | 1  | US-08-459-717-1  |
| 10         | 163.8 | 9.6         | 2075   | 1  | US-08-712-302-1  |
| 11         | 163.8 | 9.6         | 2075   | 2  | US-08-880-031-1  |
| 12         | 163.8 | 9.6         | 2075   | 3  | US-09-097-179-1  |
| 13         | 163.8 | 9.6         | 2075   | 4  | US-09-080-715-1  |
| 14         | 163.8 | 9.6         | 2075   | 4  | US-09-142-569-7  |
| 15         | 163.8 | 9.6         | 2075   | 5  | PCT-US96-08140-1 |
| 16         | 163.8 | 9.6         | 2998   | 3  | US-09-054-368-1  |
| 17         | 163.8 | 9.6         | 2998   | 3  | US-09-054-274-1  |
| 18         | 163.8 | 9.6         | 2998   | 4  | US-09-056-704-1  |
| 19         | 161   | 9.4         | 2267   | 4  | US-09-142-569-5  |
| 20         | 156.4 | 9.2         | 2350   | 4  | US-09-187-478-1  |
| 21         | 154.8 | 9.1         | 2350   | 4  | US-09-292-036-1  |
| 22         | 123.4 | 7.2         | 1418   | 4  | US-09-142-569-3  |
| 23         | 119.8 | 7.0         | 1766   | 4  | US-09-182-145-9  |
| 24         | 119.8 | 7.0         | 1766   | 4  | US-09-182-145-10 |
| 25         | 117   | 6.9         | 1480   | 4  | US-09-142-569-1  |
| 26         | 111.6 | 6.5         | 2830   | 4  | US-09-182-145-1  |
| 27         | 111.6 | 6.5         | 2830   | 4  | US-09-182-145-2  |

|    |       |     |         |   |                   |                    |
|----|-------|-----|---------|---|-------------------|--------------------|
| 28 | 102.8 | 6.0 | 1128    | 2 | US-08-459-101A-1  | Sequence 1, Appli  |
| 29 | 98.4  | 5.8 | 1062    | 4 | US-09-253-316-3   | Sequence 3, Appli  |
| 30 | 89.2  | 5.2 | 1403    | 4 | US-09-182-145-23  | Sequence 23, Appli |
| 31 | 84    | 4.9 | 4214    | 4 | US-09-122-135-1   | Sequence 1, Appli  |
| 32 | 75.4  | 4.4 | 1142    | 4 | US-09-253-316-1   | Sequence 1, Appli  |
| 33 | 75.4  | 4.4 | 1212    | 4 | US-09-182-145-34  | Sequence 34, Appli |
| 34 | 75.4  | 4.4 | 1212    | 4 | US-09-182-145-35  | Sequence 35, Appli |
| 35 | 75.4  | 4.4 | 1335    | 4 | US-09-182-145-30  | Sequence 30, Appli |
| 36 | 75.4  | 4.4 | 1335    | 4 | US-09-182-145-31  | Sequence 31, Appli |
| 37 | 65.8  | 3.9 | 1101    | 4 | US-09-182-145-29  | Sequence 29, Appli |
| 38 | 63.4  | 3.7 | 693     | 4 | US-09-182-145-24  | Sequence 24, Appli |
| 39 | 63.4  | 3.7 | 1202    | 4 | US-09-182-145-26  | Sequence 26, Appli |
| 40 | 48.6  | 2.8 | 2541    | 2 | US-08-656-393-1   | Sequence 1, Appli  |
| 41 | 41.4  | 2.4 | 51      | 4 | US-09-182-145-117 | Sequence 117, App  |
| 42 | 39.8  | 2.3 | 546     | 4 | US-09-182-145-28  | Sequence 28, Appli |
| 43 | 39.8  | 2.3 | 683     | 4 | US-09-182-145-25  | Sequence 25, Appli |
| 44 | 39.8  | 2.3 | 1183    | 4 | US-09-182-145-27  | Sequence 27, Appli |
| 45 | 36.6  | 2.1 | 4403765 | 4 | US-09-103-840A-2  | Sequence 2, Appli  |

ALIGNMENTS

RESULT 1  
US-09-182-145-17  
; Sequence 17, Application US/09182145B  
; Patent No. 6387657  
; GENERAL INFORMATION:  
; APPLICANT: Botstein, David A.  
; APPLICANT: Cohen, Robert  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth J.  
; APPLICANT: Lawrence, David A.  
; APPLICANT: Levine, Arnold J.  
; APPLICANT: Pennica, Diane  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Wood, William I.  
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: P1176R2  
; CURRENT APPLICATION NUMBER: US/09/182,145B  
; CURRENT FILING DATE: 1998-10-29  
; EARLIER APPLICATION NUMBER: US 60/063,704  
; EARLIER FILING DATE: 1997-10-29  
; EARLIER APPLICATION NUMBER: US 60/073,612  
; EARLIER FILING DATE: 1998-02-04  
; EARLIER APPLICATION NUMBER: US 60/081,695  
; EARLIER FILING DATE: 1998-04-14  
; NUMBER OF SEQ ID NOS: 156  
; SEQ ID NO 17  
; LENGTH: 1734  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-09-182-145-17

|                       |              |   |                 |                     |
|-----------------------|--------------|---|-----------------|---------------------|
| Query Match           | 74.8%;       | Score 1278;   | DB 4;           | Length 1734;        |
| Best Local Similarity | 88.5%;       | Pred. No. 0;  |                 |                     |
| Matches 1520;         | Conservative | 0;  | Mismatches 165; | Indels 33; Gaps 11; |
| QY                    | 3            | CGTCTTCGATCTCCAGAGGACCCCTGGGGTGGGACAGGGCCCTTGGCAAGGCTGCAGCGCC   | 62              |                     |
|                       |              |   |                 |                     |
| Db                    | 13           | CGTCTTCGATCTCCAGAGGACCCCGGGCTGGGACAGGGCCCTTGGCGAGGCTGCAGCTGC    | 72              |                     |
|                       |              |   |                 |                     |
| QY                    | 63           | TG-GGCAGTGGCTTGAATGGAGGCTTTATTACTGGGAACCTGAGGAGCTTAAGAGGCTCC    | 121             |                     |
|                       |              |   |                 |                     |
| Db                    | 73           | TGTGGCAGTAGCTTGGGATGGAGGCTTTCTTGTGGGAACCTGAGAGGCTCGAGAGGCTCC    | 132             |                     |
|                       |              |   |                 |                     |
| QY                    | 122          | TGTCAG---CTTGTCTCTAAAGTCTTAGCACTTGTGTGGCTGGGCTTCACACACTGTCA     | 178             |                     |
|                       |              |   |                 |                     |
| Db                    | 133          | TGTCAGGCTCTCTGTCTCTAAACTCTTTGGCACTTGGCGTGGCTTGGGCTTCACACACTGTCA | 192             |                     |
|                       |              |   |                 |                     |
| QY                    | 179          | GACACCTTCGTTGGTGGCCCTCCACGGGCTCACCTTTCAAGTTTGAAGCTGGCTCCACAAGG  | 238             |                     |
|                       |              |   |                 |                     |

|    |      |                      |                              |                           |                         |
|----|------|----------------------|------------------------------|---------------------------|-------------------------|
| Qy | 1313 | GTGCTGGGGTAAGGTC     | AAATGTTAGG-ACACAGACAGCAGATTC | CCCTGAAACTTCCAAATTC       | 1371                    |
|    |      |                      |                              |                           |                         |
| Db | 1327 | ATGCTGGGGATAGTGACAGT | AAATAGGTAC                   | CAGCAGAGATTC              | CCCTGAAACATCCAGGTC 1386 |
|    |      |                      |                              |                           |                         |
| Qy | 1372 | CTTCTTTGGGACTTCTGTAT | CTGTCCTCCCAAGATGATGAAT       | GAATCGCTAAGTGTACCT        | 1431                    |
|    |      |                      |                              |                           |                         |
| Db | 1387 | CCTTCTTTGGGACTTCTAT  | GTCTTTGT-CCCAAGATAT          | TGGGTGACCTTGTAAAGTGTGCCCT | 1445                    |
|    |      |                      |                              |                           |                         |
| Qy | 1432 | TCCTGACCTGGAACACCC   | TGCTGCTCGGGAAGTATT           | CAGGGGAGAAATTCCTGTGTGA    | 1491                    |
|    |      |                      |                              |                           |                         |
| Db | 1446 | TTCTGTATCTGGAACACCC  | TGCCGCTGGGA-----             | AGAAATTTCTTGGGA           | 1492                    |
|    |      |                      |                              |                           |                         |
| Qy | 1492 | ACATGAAGAGATGAAT     | CACACTGCTCTTAAGAAATTC        | CTTGAAGTCCAGGAACCTTGAGC   | 1550                    |
|    |      |                      |                              |                           |                         |
| Db | 1493 | ACATGAAGAGATGGAAT    | CACACTATTCTTAAGAGCGTT        | TCCCAAGTCCAGGAACCTTGACC   | 1552                    |
|    |      |                      |                              |                           |                         |
| Qy | 1551 | TTTGTATTTCAGAAAT     | CGATCTCTTAAGCACTGCC          | CAAAACAGGAGGCTCCACACCT    | 1610                    |
|    |      |                      |                              |                           |                         |
| Db | 1553 | TTTGTATTTCAGAAAT     | CACACTCTCTTAATGCT            | CCAAAGCAAG-AGGCTCCACACTT  | 1611                    |
|    |      |                      |                              |                           |                         |
| Qy | 1611 | CTGCAGCCGAGGCGCTT    | CTCTTCAGCATGAGAAGACAAG       | GACAGCAGAGTACTCTC         | 1670                    |
|    |      |                      |                              |                           |                         |
| Db | 1612 | CTGCAGCCGAGGCGCTT    | CTCTTCAGCATGAGAGACAAG        | AACAGTAGTAGTACCTC         | 1671                    |
|    |      |                      |                              |                           |                         |
| Qy | 1671 | CTCTGGAGGACTAGTCT    | TAGCTAGCAATAACACCC           | CAAA 1708                 |                         |
|    |      |                      |                              |                           |                         |
| Db | 1672 | CTCTGGAGGACTGCGC     | CGGCTGGAATAAACACCC           | CAAA 1709                 |                         |
|    |      |                      |                              |                           |                         |

RESULT 2  
 US-09-182-145-18/c  
 ; Sequence 18, Application 05/09182145B  
 ; Patent No. 6387657  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Botstein, David A.  
 ; APPLICANT: Cohen, Robert  
 ; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Gurney, Austin L.  
 ; APPLICANT: Hillan, Kenneth J.  
 ; APPLICANT: Lawrence, David A.  
 ; APPLICANT: Levine, Arnold J.  
 ; APPLICANT: Pennica, Diane  
 ; APPLICANT: Roy, Margaret Ann  
 ; APPLICANT: Wood, William F.  
 ; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME  
 ; FILE REFERENCE: P1176R2  
 ; CURRENT APPLICATION NUMBER: US/09/182,145B  
 ; CURRENT FILING DATE: 1998-10-29  
 ; EARLIER APPLICATION NUMBER: US 60/063,704  
 ; EARLIER FILING DATE: 1997-10-29  
 ; EARLIER APPLICATION NUMBER: US 60/073,612  
 ; EARLIER FILING DATE: 1998-02-04  
 ; EARLIER APPLICATION NUMBER: US 60/081,695  
 ; EARLIER FILING DATE: 1998-04-14  
 ; NUMBER OF SEQ ID NOS: 156  
 ; SEQ ID NO 18  
 ; LENGTH: 1734  
 ; TYPE: DNA  
 ; ORGANISM: Mus musculus  
 US-09-182-145-18

|                       |                 |                 |            |              |
|-----------------------|-----------------|-----------------|------------|--------------|
| Query Match           | 74.88;          | Score 1278;     | DB 4;      | Length 1734; |
| Best Local Similarity | 88.58;          | Pred. No. 0;    |            |              |
| Matches 1520;         | Conservative 0; | Mismatches 165; | Indels 33; | Gaps 11;     |

|    |      |                        |                          |                              |      |
|----|------|------------------------|--------------------------|------------------------------|------|
| Qy | 3    | CGCTTCTGTATCTCCAGAGGAC | CCCTGGGCTGGGACAGGGCGCTT  | TGGCAAGGCTGCAGCGCGC          | 62   |
|    |      |                        |                          |                              |      |
| Db | 1722 | CGTCTCTGATCTCCAGAGG    | ACCCCGGCTGGGACAGGGGCGCTT | TGGCAGGCGCTGCAGCTGC          | 1663 |
|    |      |                        |                          |                              |      |
| Qy | 63   | TG-GGCAGTGGCTTGG       | GAATCGAGGTCTTTTATTACT    | TGGGAACCTGAGGAGCTTAAGAGGCTCC | 121  |
|    |      |                        |                          |                              |      |
| Db | 1662 | TGTGGCAGTAGCTTGG       | ATGGATCGAGTCTTTCTTGT     | TGGGAACCTGAGGAGCTGAGAGGCTCC  | 1603 |
|    |      |                        |                          |                              |      |

RESULT 2  
US-09-182-145-18/c  
; Sequence 18, Application US/09182145B

```

: GENERAL INFORMATION:
: APPLICANT: Botstein, David A.
: APPLICANT: Cohen, Robert
: APPLICANT: Goddard, Audrey
: APPLICANT: Gurney, Austin L.
: APPLICANT: Hillan, Kenneth J.
: APPLICANT: Lawrence, David A.
: APPLICANT: Levine, Arnold J.
: APPLICANT: Pennica, Dianek
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Wood, William T.
: TITLE OF INVENTION: WISP POLY
: FILE REFERENCE: P1176R2
: CURRENT APPLICATION NUMBER: U
: CURRENT FILING DATE: 1998/10
: EARLIER APPLICATION NUMBER: U
: EARLIER FILING DATE: 1997/10
: EARLIER APPLICATION NUMBER: U
: EARLIER FILING DATE: 1998/02
: EARLIER APPLICATION NUMBER: U
: EARLIER FILING DATE: 1998/04
: NUMBER OF SEQ ID NOS: 156
: SEQ ID NO 18
: LENGTH: 1734
: TYPE: DNA
: ORGANISM: Mus musculus
: US-09-182-145-18

```

[illegible]

QY 122 TGTCTGAG---CTTGTCTCTAAAGTCTTAGCACCTTGTGGTGGCTTGCACACACTGTCA 178  
DB 1602 TGTCTGAGTCTGTCTCTAAAGTCTTAGCACCTTGTGGTGGCTTGCACACACTGTCA 1543  
QY 179 GACACCTTGTGTGGCTTCCAGGCGCTCACCTTCAAGTTTGAAGTGGCTTCCACAAGGG 238  
DB 1542 GACACCTTGTGTGGCTTCCAGGCGCTCACCTTCAAGTTTGAAGTGGCTTCCACAAGGG 1489  
QY 239 ACAGGTGACATGAGGGGAGCCACTGATCCATCTTCTGGCCACTTCTCTCTCTGCT 298  
DB 1488 ACAGGTGACATGAGGGGAGCCACTGATCCATCTTCTGGCCACTTCTCTCTCTGCT 1429  
QY 299 TCTCTCAATGGTGTGCCAGCTGTCCGGGACACCTGTGATCTGCTTGTGACACCAACC 358  
DB 1428 TCTCTCAATGGTGTGCCAGCTGTCCGGGAGCCACTGTGATCTGCTTGTGACACCAACC 1369  
QY 359 CAGTGTCCACAGGGGATACCCCTGTGTGGTGGCTGTGATCTGCTTGTGATCTGCTG 418  
DB 1368 CAGTGTCCACAGGGGATACCCCTGTGTGGTGGCTGTGATCTGCTTGTGATCTGCTG 1309  
QY 419 ACAGGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTG 478  
DB 1308 ACAGGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTG 1249  
QY 479 TGTCTAGCCTTGGGGGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTG 538  
DB 1248 TGTCTAGCCTTGGGGGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTG 1189  
QY 539 CGTGTAGCTGTAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTG 598  
DB 1188 CGGAGCTGTAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTG 1129  
QY 599 CAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTG 658  
DB 1128 CAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTG 1069  
QY 659 TGTGCGGCTGCGGAGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTG 718  
DB 1068 TGTGCGGCTGCGGAGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTG 1009  
QY 719 CTGCCCCGAGTGGTATGTGACAGGAGTGA---CACGGGAGTCCAGGCTCCAGGC 775  
DB 1008 CTGCCCCGAGTGGTATGTGACAGGAGTGA---CACGGGAGTCCAGGCTCCAGGC 949  
QY 776 GCAAGGACACCACTTCTGCGCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 835  
DB 948 CCAAGGACACCACTTCTGCGCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 889  
QY 836 TTGGAGCAGAGCTGGGGGCGCTGTCTCAACCACTGTGGCTGGGATAGCCACCGAGT 895  
DB 888 CTGGAGCAGAGCTGGGGGCGCTGTCTCAACCACTGTGGCTGGGATAGCCACCGAGT 829  
QY 896 GTCCACCAAGACCGATTTGCCAAGTGGAGTCCAGGCTGTGCTGTGCTGTGCTGTGCT 955  
DB 828 ATCCACCAAGACCGATTTGCCAAGTGGAGTCCAGGCTGTGCTGTGCTGTGCTGTGCT 769  
QY 956 CTGCTGGGAGCAGGAGCAGCAGCTCATGGAACAGTGTCTTCAAGGCGCA-ACCTGGGA 1014  
DB 768 CTGCTGGGATCCAGGAGCAGGCTCATGGAACAGTGTCTTCAAGGCGCA-ACCTGGGA 709  
QY 1015 TGGGGATACAGGCGCTGCCATCTCAGCAAAATGAGCCATAGGACCGGCTGTGAGTGTG 1074  
DB 708 TGTGGATACAGGCGCTGCCATCTCAGCAAAATGAGCCATAGGACCGGCTGTGAGTGTG 649  
QY 1075 GTAGATGCT 1134  
DB 648 GTAGATGCT 589  
QY 1135 AGAGCCACTGAGGAGTCCCTGTCTCTGAGGTAGGCGGAGGAGTGTGAGGCTCCAGTCCAGT 1194  
DB 588 AGAGCCCTGAGGAGTCCCTGTCTCTGAGGTAGGCGGAGGAGTGTGAGGCTCCAGTCCAGT 529

QY 1195 TCTCTGG--TTCAGCCTGGAATTTCTGGGTTCTCTGGCTCATTCCTCAAAACATCCCTGT 1252  
DB 528 TCTCTGGATTTCTGACCCAGGCTTCTGGGTTCTCTGGCTAGTTCTCTCAAAACATTCCTGT 469  
QY 1253 ACAAAAGACACACCAAAAGACCTTTAAACCTTAGCTATGCTGGGCAAACTGGCCACC 1312  
DB 468 ATGAAAGACACCAAAAGACCTTTAAAGCTTAAGCTGTACTGGGCAAGCCTGGCCACC 409  
QY 1313 GTGCTGGGATAGGTTCAATGTTAGG-ACCAGACACAGATTCGCTGAAACTTCCAAATTC 1371  
DB 408 ATGCTGGGATAGTACAGTATAGGTACAGCAGCAGATTCGCTGAAACATCCAGGTC 349  
QY 1372 CTTCTTGGACTTCTGTATGCTTGTCCCAAGATGATGAATGAACCTCGTAAGTGTACCT 1431  
DB 348 CTTCTTGGACTTCTGTATGCTGTG-CCCAAGATTTAGGTTGACCTTGTAACTGTGCT 290  
QY 1432 TCCCTGACCTGAGAACACCTGCTGCTCGGGAAGTATTCAGGGGGCAGATTCCTCTGTGA 1491  
DB 289 TTTCTGATCTGAGAACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 243  
QY 1492 ACATGAAGAGAT-GAATCACACTGTCTTAAAGAAATTCCTGAAAGTCCAGGAATTCGAGC 1550  
DB 242 ACATGAAGAGATGGAATCACACTATTTCTTAAGAGCGTTTCCCAAGTCCAGGAATTCGAGC 183  
QY 1551 TTTGTATTTTCAAGGAATGACATCTCTTAAGCACTCGCAAAACAGGAAGCTCCACACT 1610  
DB 182 TTTGTATTTTAAAAATACACATCTCTTAAATGCTCAAAAGCAAG-AGGCTCCACACTT 124  
QY 1611 CTGGCAGGCGGCGCTTCTCTTTCAGCATGAGAAAGACAGGAGCAGAGTACTCTC 1670  
DB 123 CTGGCAGGCGGCGCTTCTCTTTCAGCATGAGAGAGACAGGAGCAGAGTACTCTC 64  
QY 1671 CTCTGAGGACTAGTCTAGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCT 1708  
DB 63 CTCTGAGGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 26

## RESULT 3

US-09-182-145-13  
; Sequence 13, Application US/09182145B  
; Patent No. 6387657  
; GENERAL INFORMATION:  
; APPLICANT: Botstein, David A.  
; APPLICANT: Cohen, Robert  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth J.  
; APPLICANT: Lawrence, David A.  
; APPLICANT: Levine, Arnold J.  
; APPLICANT: Pennica, Diane  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Wood, William I.  
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: P1176R2  
; CURRENT APPLICATION NUMBER: US/09/182,145B  
; CURRENT FILING DATE: 1998-10-29  
; EARLIER APPLICATION NUMBER: US 60/063,704  
; EARLIER FILING DATE: 1997-10-29  
; EARLIER APPLICATION NUMBER: US 60/073,612  
; EARLIER FILING DATE: 1998-02-04  
; EARLIER APPLICATION NUMBER: US 60/081,695  
; EARLIER FILING DATE: 1998-04-14  
; NUMBER OF SEQ ID NOS: 156  
; SEQ ID NO 13  
; LENGTH: 1293  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-182-145-13

Query Match 32.9%; Score 561.4; DB 4; Length 1293;  
Best Local Similarity 78.0%; Pred. No 2.2e-152;  
Matches 701; Conservative 0; Mismatches 196; Indels 2; Gaps 2;

```

QY 243 GGTGACATGAGGGGAGCCCACTGATCCATCTTCTGGCCACTTCTCTCTCTGCTTCTC 302
Db 16 GGGGACATGAGGAGGAGCCCAAGAACCCACCTCTCTCTCTCTCTCTCTCTCTCTCTC 75
QY 303 TCAATGTTGTTGGCCGACCTGTCGGGACCCCTGTACTCTCTCTCTCTCTCTCTCTCT 362
Db 76 TCAAGGTGGCTACCCAGCTGTGCCGACACCATGTACTCTCTCTCTCTCTCTCTCTCT 135
QY 363 TGGCCACAGGGGATACCCCTGTGTGATGGCTGTGGCTGTCTGTAAAGTGTGTGCAGG 422
Db 136 TGGCCGCTGGGAGTACCCCTGTGTGATGGCTGTGGCTGTCTGTAAAGTGTGTGCAGG 195
QY 423 AGGCTGGGGAGTCTGGGACCACTGTGATCTGTGCGACCCCTGTGCGAGGCTGTGTTGT 482
Db 196 CGGCTGGGGAGCCCTGGCGACCACTTCCACGCTCTGCGACGCGAGGCTGTGTTGT 255
QY 483 CAGCTGGGGAGCCCTGGCGGCGCATGGGCTGTGTCTCTCTCTCTCTCTCTCTCTCT 542
Db 256 CAGCCGGGGAGGAGCCCGTGGCGGGGGCCCTGTGCTCTCTCTCTCTCTCTCTCTCT 315
QY 543 AGCTGTGAGGTGAATGGCGGAGGTACCTGGATGGAGAGCTTTAAACCCCAATTCAGG 602
Db 316 AGCTGTGAGGTGAACGGCGGCTGTATCGGGAAGGGAGACCTTCCAGCCCACTGCAGC 375
QY 603 GTCCTGTCCGCTGTGATGACGGTGGCTTACCTGCTGCGCTGCTGTGCGAGTGTG 662
Db 376 ATCCGCTCCGCTGTGAGGAGGGGGCTTCACTTCTGCGCTGTGCGAGGAGTGTG 435
QY 663 CGGCTGCCAGCTGGGACTGCCAGCCGCGGAGAGATACAGGTGCCAGGAGTGTGCTG 722
Db 436 CGGCTGCCAGCTGGGACTGCCAGCCGCGGAGGAGTGTGCGCTGTGCGAGTGTGCTG 495
QY 723 CCCGAGTGGGTATGTGAGGAGGTGACACCGCGGATGCCAGCGCTTCCAGCGCGCAAGGA 782
Db 496 CCTGAGTGGGTGTGCGGCAAGAGGGGAGTGGGGACCCAGCCCTTCCAGCGCGCAAGGA 555
QY 783 CACCAACTTTCTGCCCTTGTGACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 842
Db 556 CCCAGTTTCTTGCCCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 615
QY 843 ACAGCTGGGGCCCTGCTCAACACCTGTGGCTGGGCTGGGCTAGCCACCCAGTGTCCAA 902
Db 616 AGGGCTGGGAGCCCTGCTGACCACTGTGGCTGGGCTGGGCTAGCCACCCAGTGTCCAA 675
QY 903 CAGAACCGATTCTGCCACTGGAGATCCAAAGCCGCTGTGTCTGCTGCGAGCCCTGCTG 962
Db 676 CAGAACCGCTTCTGCCAGTGGAGACCCAGCCAGCCGCTGTGTCTGCTGCGAGCCCTG 735
QY 963 GCAGCAGGAGCCAGCTCATGAAACAGTGTCTTCTA - AGGCCAACTGGGGATGCGGAT 1021
Db 736 CCCTCCAGGGTTCGAGTCCCAAAACAGTGTCTTCTAGAGCCGGGCTGGGAATGGGAC 795
QY 1022 ACAGGGCTGCTATCTCAGCAATAGCCCTAGGACAGGCGCTGGGCTGTGTGTAGATG 1081
Db 796 AGGTGTCCACCTATCCAGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGT 855
QY 1082 CT-CTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1139
Db 856 GTCCGTGCCAGGGCCCTTGGCTGCGAGGCAACACTTACCTTGGGTGCCACCATGCA 914

```

RESULT 4

US-09-182-145-14/c  
; Sequence 14, Application US/09182145B  
; Patent No. 6387657  
; GENERAL INFORMATION:  
; APPLICANT: Botstein, David A.  
; APPLICANT: Cohen, Robert  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth J.  
; APPLICANT: Lawrence, David A.  
; APPLICANT: Levine, Arnold J.

```

; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P117682
; CURRENT APPLICATION NUMBER: US/09/182,145B
; CURRENT FILING DATE: 1998-10-29
; EARLIER APPLICATION NUMBER: US 60/063,704
; EARLIER FILING DATE: 1997-10-29
; EARLIER APPLICATION NUMBER: US 60/073,612
; EARLIER FILING DATE: 1998-02-04
; EARLIER APPLICATION NUMBER: US 60/081,695
; EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 14
; LENGTH: 1293
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-182-145-14

Query Match      32.9%; Score 561.4; DB 4; Length 1293;
Best Local Similarity 78.0%; Pred. No. 2.2e-152;
Matches 701; Conservative 0; Mismatches 196; Indels 2; Gaps 2;

QY 243 GGTGACATGAGGGGAGCCCACTGATCCATCTTCTGGCCACTTCTCTCTCTGCTTCTC 302
Db 1278 GGGGACATGAGGAGGAGCCACACCAAGAACCCCTCTCTCTCTCTCTCTCTCTCTCTC 1219
QY 303 TCAATGTTGTTGCCAGCTGTGCGGACACCCCTGTACTCTCTCTCTCTCTCTCTCTCT 362
Db 1218 TCAAGGTGGCTATCCAGCTGTGCCGACACCATGTACTCTCTCTCTCTCTCTCTCTCT 1159
QY 363 TGGCCACAGGGGATACCCCTGTGTGATGGCTGTGGCTGTCTGTAAAGTGTGTGCAGG 422
Db 1158 TGGCCGCTGGGAGTACCCCTGTGTGATGGCTGTGGCTGTCTGCGGGTATGTGCAGG 1099
QY 423 AGGCTGGGGAGTGTCTGCGACCACTGTCTGCGACCCAGCCAGCGCTGTGTTGT 482
Db 1098 CGGCTGGGGAGCTGTCGACCACTTCCACGCTCTGCGACGCGCAGCAGGCTGTCTG 1039
QY 483 CAGCTGGGGAGCCCTGCGGCGCATGGGCTGTGTCTCTCTCTCTCTCTCTCTCTCTCT 542
Db 1038 CAGCCCGGGGAGGAGCCCGTGGCGGGGGCCCTGTGCTCTCTCTCTCTCTCTCTCTCT 979
QY 543 AGCTGTGAGGTGAATGGCGGAGGTACCTGGATGGAGAGACCTTTAAACCCCAATTCAGG 602
Db 978 AGCTGTGAGGTGAACGGCGGCTGTATCGGGAAGGGAGACCTTCCAGCCCACTGCAGC 919
QY 603 GTCCTGTCCGCTGTGATGACGGTGGCTTACCTGCTCTGCGCTGTGCGAGTGTG 662
Db 918 ATCCGCTGCCCTGTGGAGGAGGGGCTTCACTCTGCTGCGCTGTGCGAGGAGTGTG 859
QY 663 CGGCTGCCAGCTGGGACTGCCAGCCGCGGAGAGATACAGGTGCCAGGAAAGTGTG 722
Db 858 CGGCTGCCAGCTGGGACTGCCAGCCGCGGAGGAGGTGAGGAGTCTTGGGCAAGTGTG 799
QY 723 CCCGAGTGGGTATGTGACAGGAGTGCACAGCGGCGATGCCAGCGCTTCCAGCGCGCAAGGA 782
Db 798 CCTGAGTGGGTGTGGGCAAGAGGGGAGTGGGAGCCCAAGCCCTTCCAGCCCAAGGA 739
QY 783 CACCAACTTCTGCTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 842
Db 738 CCCAGTTTCTTGCGCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 679
QY 843 ACAGCTGGGGCCCTGCTCAACACCTGTGGCTGGGCTAGCATAGCCACCGAGTGTCCAA 902
Db 678 ACGGCTGGGGAGCCCTGCTCAACCACTGTGGCTGGGCTAGCATAGCCACCGAGTGTCCAA 619
QY 903 CAGAACCGATTCTGCGCAACTGGAGATCCAAAGCCGCTGTGTCTCTCTCTCTCTCTCT 962
Db 618 CAGAACCGCTTCTGCGGAGTGTGAGACCCAGCGCGCTGTGCTCTCTCTCTCTCTCTCT 559
QY 963 GCAGCAGGAGCCCAAGCTCATGAGAACAGTGTCTTCTA - AGGCCAACTGGGGATGCGGAT 1021

```

Db 558 CCCTCCAGGGTCCAGTCCACAAACAGTGCCTTTAGAGCCGGGCTGGGAATGGGAC 499  
Qy 1022 ACAGGGCTGCATCCCTCAGCAATGACCCCTAGACACAGGCCCTGGACTGCTGTAGATG 1081  
Db 498 ACGGTGTCCACCATCCCGAGTGTGGCCCTGTGCCCTGGGCCCTGGGCTGATGAAGATG 439  
Qy 1082 CT-CTTCCATGCTTGTGGCTGCGAGTTAACTGTCTCTGCTTGGATTCACTGTGTAGAGC 1139  
Db 438 GTCCGTGCCAGGCCCTTGGCTGCAGGCAACACATTTAGCTTGGTGTCCACCATGCAGAAC 380

## RESULT 5

US-09-182-145-39

; Sequence 39, Application US/09182145B

; Patent No. 6387657

; GENERAL INFORMATION:

; APPLICANT: Botstein, David A.

; APPLICANT: Cohen, Robert

; APPLICANT: Goddard, Audrey

; APPLICANT: Gurney, Austin L.

; APPLICANT: Hillan, Kenneth J.

; APPLICANT: Lawrence, David A.

; APPLICANT: Levine, Arnold J.

; APPLICANT: Pennica, Diane

; APPLICANT: Roy, Margaret Ann

; APPLICANT: Wood, William I.

; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME

; FILE REFERENCE: P117682

; CURRENT APPLICATION NUMBER: US/09/182,145B

; CURRENT FILING DATE: 1998-10-29

; EARLIER APPLICATION NUMBER: US 60/063,704

; EARLIER FILING DATE: 1997-10-29

; EARLIER APPLICATION NUMBER: US 60/073,612

; EARLIER FILING DATE: 1998-02-04

; EARLIER APPLICATION NUMBER: US 60/081,695

; EARLIER FILING DATE: 1998-04-14

; NUMBER OF SEQ ID NOS: 156

; SEQ ID NO 39

; LENGTH: 841

; TYPE: DNA

; ORGANISM: Artificial sequence

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: 1-841

; OTHER INFORMATION: Sequence is synthesized.

; Patent No. 6387657

US-09-182-145-39

Query Match

Best Local Similarity 31.0%; Score 528.8; DB 4; Length 841;

Matches 658; Conservative 0; Mismatches 177; Indels 3; Gaps 2;

Qy 243 GGTGACATGAGGGGACGCCCACTGATCCATCTTGTGCCACTTCTTCTCTGCTGCTCTC 302  
Db 6 GGGGACATGAGGACGACACCAACCCACCTCTGCGCTTCTCCCTCTCTGCTCTCTC 65  
Qy 303 TCAATGTTGTGCGCCAGCTGTGCGGACACCCCTGTACCTGTCTTGGACACCAACCCAG 362  
Db 66 TCAAGTGTGCTACCCAGCTGTGCGGACACCACTGTACCTGCGCCCTGGCCACCTCCCGA 125  
Qy 363 TGCCACAGGGGGTACCCCTGGTGTGCTGATGGCTGTGGCTGCTCTTAAAGTGTGTACGG 422  
Db 126 TGCCCGCTGGAGTACCCCTGGTGTGATGGCTGTGGCTGCTCTGCGGGTGTGTGCACGG 185  
Qy 423 AGGTGTGGGAGTCTCGGACCACTGCTGATGTCTGACACCCGACCGAGGGGCTGGTTGT 482  
Db 186 CGGTGTGGGAGGCTCGGACCACTGCTGATGTCTGACACCCGACCGAGGGGCTGGTCTGC 245  
Qy 483 CAGCTGTGGGAGGCTCGGAGGCTGATGGGCTGTGTCTTGGATGAGGATGACGGT 542  
Db 246 CAGCCCGGGAGGACCCCGTGGCGGGGGGCGCCCTGTGCTCTTGGCAGAGGAGCAGCAGC 305



Qy 836 TTGAGCAGACGCTGGGGCCCTCTCAACACCTGTGGGCTGGGCATAGCCACCCGAGT 895  
Db 732 CCAGCACCAGAGTGAGCGCCTGTTCAGACCTGTGGATGGCATCTCCACCCGGT 791  
Qy 896 GTCCAACAGACGATTCGCCAACTGGAGATCCACGCGCGCTGTCTGCCACACC 955  
Db 792 TACCAATGACAACGCCCTCTGACGCTAGAGAGCAGAGCGCGCTGTGCATGTCAGGCC 851  
Qy 956 CTGCCTGGCAGCC 968  
Db 852 TTGCGAAGCTGAC 864

RESULT 8  
US-08-386-680-1  
; Sequence 1, Application US/08386680  
; Patent No. 5585270  
GENERAL INFORMATION:  
; APPLICANT: Grotendorst, Gary R.  
; APPLICANT: Bradham Jr., Douglas M.,  
; TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Spensley Horn Jubas & Lubitz  
; STREET: 4225 Executive Square, Suite 1400  
; CITY: La Jolla  
; STATE: CA  
; COUNTRY: US  
; ZIP: 92037  
COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/386,680  
; FILING DATE: 10-FEB-1995  
CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/167,628  
; FILING DATE:  
; APPLICATION NUMBER: US/07/752,427  
; FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
; NAME: Wetherell, Jr. Ph.D., John W.  
; REGISTRATION NUMBER: 31,678  
; REFERENCE/DOCKET NUMBER: PD-1294  
TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619-455-5100  
; TELEFAX: 619-455-5110  
; INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
; LENGTH: 2075 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
IMMEDIATE SOURCE:  
; CLONE: DB60R32  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 130..1177  
US-08-386-680-1

Query Match 9.68; Score 163.8; DB 1; Length 2075;  
Best Local Similarity 53.5%; Pred. No. 2.1e-37;  
Matches 392; Conservative 0; Mismatches 332; Indels 9; Gaps 2;

Qy 242 CGGTGACATGAGGGGAGCCACATGATCTCTGGCCACTTCTCTCTCTCTCTCTCT 301  
Db 135 CGCGCCAGTATGGGCGCCGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 194  
Qy 302 CTCAATGGTGTGTGCCCGAGCTGTGCCGACACCCCTGTACCTGTCTTGGACACCA 361

Db 195 GCCGGCGTGGCGCAGAACTGCAGCGCGCGCTGCCGTGCCCGACGAGCGCGCGCG 254  
Qy 362 GTGCCACACAGGGGTACCCCTGGTGTGGATGGCTGTGGCTGTGAAGTGTGTGCACG 421  
Db 255 CTGCCCGCGGGGTGAGCCTCGTGTGGAGCGGTGCGGTGTGCGCGCTCTCGCGCAA 314  
Qy 422 GAGCTGGGGAGTCCCTGCGACCACTGTCATGTCTGCGACCCCGACGAGCGCCCTGTTG 481  
Db 315 GCAGCTGGCGAGCTGTGCACCGAGCGCGACCCCTGCGACCCCGACAGGSCCTCTTCTG 374  
Qy 482 TCAGCTGGGGAGGCCCTGGCGGCATGGGGCTGTGTCTTGTGATAGGATGACGG 541  
Db 375 TGACTTGGCTCCCGGCCAACCGCAAGATGCGCGTGTG---CACCGCCCAAGATGTTGC 431  
Qy 542 TAGCTGTGAGTGAATGCGCGCAGGTACCTGGATGGAGACACCTTTAAACCAATTCAG 601  
Db 432 TCCCTGCATCTTGGTGTACGGTGTACCGAGCGGAGATCTCTCCAGAGCACTGCAA 491  
Qy 602 GGTCTGTGCGCTGTGATGACGTGGCTTTCACCTGCTGCCGTGTGCGATGAGGATGT 661  
Db 492 GTACCACTGCACGTGCGTGGACGGGGGTGGCTGCATGCCCTGTGACGATGGAGT 551  
Qy 662 CGGCTCCCGAGCTGGGACTGCCACCGCCCAAGAGAAATACAGTGGCAGGAAAGTGTG 721  
Db 552 TCCTGTGCCAGCCCTGACTGCCCTTCCGAGGAGGGTCAAGCTGCCCGGAAATGCTG 611  
Qy 722 CCGGAGTGGTATGACCAGGGAGTGACACCGGATCCAGGCGCTCCAGCGCGCAAGG 781  
Db 612 CGAGGAGTGGTGTGACGAGCCCAAGAGCAACCGCTGGTGGGCTGCGCCCTCGCGGC 671  
Qy 782 ACACCAACTTCTGCCCTTGTCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 835  
Db 672 TTACCGACTGGAAGACACGTTTGGCCCGACAGCCCACTATGATTAGACCACTGCTG 731  
Qy 836 TTGAGCAGACGCTGGGGCCCTGCTCAACACCTGTGGGCTGGGCATAGCCACCCGAGT 895  
Db 732 CCAGACCACAGTGGAGCGCTGTTCCAAAGACCTGTGGATGGGCATCTCCACCCGGT 791  
Qy 896 GTCCAACCAAGACCGATTCGCCAACTGGAGATCCAAAGCGCGCTGTGCTGCCACACC 955  
Db 792 TACCAATGACAACGCCCTCTGACGCTAGAGAGCAGAGCGCGCTGTGTCATGGTCAGGCC 851  
Qy 956 CTGCCTGGCAGCC 968  
Db 852 TTGCGAAGCTGAC 864

RESULT 9  
US-08-459-717-1  
; Sequence 1, Application US/08459717  
; Patent No. 5770209  
GENERAL INFORMATION:  
; APPLICANT: Grotendorst, Gary R.  
; APPLICANT: Bradham Jr., Douglas M.,  
; TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Spensley Horn Jubas & Lubitz  
; STREET: 4225 Executive Square, Suite 1400  
; CITY: La Jolla  
; STATE: CA  
; COUNTRY: US  
; ZIP: 92037  
COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/459,717  
; FILING DATE: 02-JUN-1995  
; CLASSIFICATION: 536

[illegible]







NAME: Wetherell, Jr. Ph.D., John W.

REGISTRATION NUMBER: 31,678

REFERENCE/DOCKET NUMBER: PD-1294

TELECOMMUNICATION INFORMATION:

TELEPHONE: 619-455-5100

TELEFAX: 619-455-5110

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 2075 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

IMMEDIATE SOURCE:

CLONE: DB60R32

FEATURE:

NAME/KEY: CDS

LOCATION: 130..1177

US-09-080-715-1

Query Match 9.6%; Score 163.8; DB 4; Length 2075;  
Best Local Similarity 53.5%; Pred. No. 2.1e-37;  
Matches 392; Conservative 0; Mismatches 332; Indels 9; Gaps 2;

```
QY 242 CGGTGACATGAGGGGACCCCACTGATCCATCTTGTGCCACTTCTTCCCTCTGCTTCT 301
Db 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
135 GCGCGCCAGTATGGCCCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCG 194
QY 302 CTCATGCTGTGTCACAGCTGTGCCGACACCTGTACCTGTCTTGGACACACCCCA 361
Db 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
195 GCGGCGCTCGCCAGAACTGACGGGCGCTGCGGTCGCGGACGACGCGGCGCGCG 254
QY 362 GTGCCACAGGGGTACCCCTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 421
Db 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
255 CTGCGCGCGCGCTGAGCCCTGCTGCTGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 314
QY 422 GAGCGTGGGAGGCTGCGACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 481
Db 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
315 GCAGCTGGCGAGCTGTGACCGCGCGACCCCTGCGACCGCGACCGCGCGCTTCTG 374
QY 482 TCAGCGTGGGCGCGCTGCGCGCATGCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 541
Db 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
375 TGACTTCGGCTCCCGGCCAACCGCAAGATCGGCGTGTG---CACCGCAAGATGGTGC 431
QY 542 TAGCTGTGAGTGAATGGCCGACGATGCTGATGGAGAGACCTTTAAACCAATTCGAG 601
Db 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
432 TCCCTGCATCTTCGTGTGTACGGTGTACCGAGCGGAGAGTCTTCCAGACGAGCTGCAA 491
QY 602 GGTCTGTGCGCTGTGATGACGCTGTGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTG 661
Db 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
492 GTACCACTGACGCTGCTGAGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 551
QY 662 GCGCTGCCAGCTGGGACTGCCACGCCCAAGAGATACAGGTGCGCAGGAAAGTGTG 721
Db 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
552 TCCTGTGCGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 611
QY 722 CCGCAGTGGGTATGTACCAAGGAGTGACACCGCGCATCCAGCGCTCCAGCGGCGCAAG 781
Db 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
612 CGAGAGTGGGTGTGTGACAGAGCCCAAGCAACCGTGTGCGCTGCGCTGCGCGC 671
QY 782 ACACCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 835
Db 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
672 TTACCGATGGAAGACAGCTTTGGCCCGAGACCACTATGATAGAGCAACTGCTGCTG 731
QY 836 TTGAGCACACGCTGGGCGCTGCTCAACACCTGTGGCTGGGCGATAGCACCCCGAGT 895
Db 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
732 CCAGACACAGTGGAGCGCTGTTCACAGACCTGTGGGATGGGATCTCCACCCCGGGT 791
QY 896 GTCCAAACAGAACGATTTCTGCCAACTGGAGTCCAAACCGCGCTGTGCTGCGCCAGACC 955
Db 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
792 TACCAATGACAACGCTCTCTGACGCTAGAGAAGCAGACGCGCTGTGCTGCTGCTGCTG 851
QY 956 CTGCTTGGCAGCC 968
```

```
Db 852 TTGCGAAGCTGAC 864
111 111 1
RESULT 14
US-09-142-569-7
; Sequence 7, Application US/09142569
; Patent No. 6413735
; GENERAL INFORMATION:
; APPLICANT: Lau, Lester F.
; TITLE OF INVENTION: Extracellular Matrix Signalling Molecules
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 02-Apr-1999
; CLASSIFICATION: <unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Clough, David W.
; REGISTRATION NUMBER: 36,107
; REFERENCE/DOCKET NUMBER: 28758/33766
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2075 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: "CTGF cDNA coding sequence"
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-142-569-7
```

```
Query Match 9.6%; Score 163.8; DB 4; Length 2075;
Best Local Similarity 53.5%; Pred. No. 2.1e-37;
Matches 392; Conservative 0; Mismatches 332; Indels 9; Gaps 2;
QY 242 CGGTGACATGAGGGGACCCCACTGATCCATCTTGTGCCACTTCTTCCCTCTGCTTCT 301
Db 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
135 GCGCGCCAGTATGGCCCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCG 194
QY 302 CTCATGCTGTGTCACAGCTGTGCCGACACCTGTACCTGTCTTGGACACACCCCA 361
Db 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
195 GCGGCGCTCGCCAGAACTGACGGGCGCTGCGGTCGCGGACGACGCGGCGCGCG 254
QY 362 GTGCCACAGGGGTACCCCTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 421
Db 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
255 CTGCGCGCGCGCTGAGCCCTGCTGCTGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 314
QY 422 GAGCGTGGGAGGCTGCGACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 481
Db 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
315 GCAGCTGGCGAGCTGTGACCGCGCGACCCCTGCGACCGCGACCGCGCGCTTCTG 374
QY 482 TCAGCGTGGGCGCGCTGCGCGCATGCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 541
Db 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
375 TGACTTCGGCTCCCGGCCAACCGCAAGATCGGCGTGTG---CACCGCAAGATGGTGC 431
```

QY 542 TAGCTGTAGGTGAATGGCCGAGGTACCTGTGATGGAGAGACCTTTAAACCCAAATTCGAG 601  
Db 432 TCCCTGCTACTTCGGTGTGACGGTGTACCCGACGCGAGAGTCTCTCCAGAGCAGCTGCAA 491  
QY 602 GGTCTGTGCCCTGTGATGACGGTGTACCTGCGCTGCGCTGTGAGTGAAGTGT 661  
Db 492 GTACCAAGTCCAGTGTGATGACGGTGTGCGCTGTGAGTGTGAGTGTGAGTGTGAGTGT 551  
QY 662 GCGGCTGCCAGTGTGATGACGGTGTGCGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 721  
Db 552 TCGTGTGCCAGCTGTGATGACGGTGTGCGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 611  
QY 722 CCCCAGTGGGTATGTGACCAAGGAGTGTGACCGGGGATCCAGGCTCCACGGCGCAAGG 781  
Db 612 CGAGGAGTGGGTGTGTGACGAGCCCAAGGACCAACCGTGTGGGCTTGCCTCGCGGC 671  
QY 782 ACACCAACTTCTGCCCTGTGATGACGGTGTGCGCTGTGAGTGTGAGTGTGAGTGTGAGTGT 835  
Db 672 TTACCGACTGGAGACAGCTTTGGCCGACGACCAACTATGATAGAGCCAACTGCCGTGT 731  
QY 836 TTGGAGCAGAGCTGTGATGACGGTGTGCGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 895  
Db 732 CCAGACCAAGAGTGTGATGACGGTGTGCGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 791  
QY 896 GTCCACCAAGAGTGTGATGACGGTGTGCGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 955  
Db 792 TACCAATGACAGCTGTGATGACGGTGTGCGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 851  
QY 956 CTGCTGCGAGCC 968  
Db 852 TTGCGAAGCTGAC 864

## RESULT 15

PCT-US96-08140-1  
; Sequence 1, Application PC/TUS9608140  
; GENERAL INFORMATION:  
; APPLICANT: University of South Florida  
; TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FISH & RICHARDSON P.C.  
; STREET: 4225 Executive Square, Suite 1400  
; CITY: La Jolla  
; STATE: CA  
; COUNTRY: US  
; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US96/08140  
; FILING DATE: 30-MAY-1996  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Haile, Ph.D., Lisa A.  
; REGISTRATION NUMBER: 38,347  
; REFERENCE/DOCKET NUMBER: 07414/003W01  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619-678-5070  
; TELEFAX: 619-678-5099  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2075 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; IMMEDIATE SOURCE:  
; CLONE: CTGF  
; FEATURE:

; NAME/KEY: CDS  
; LOCATION: 130...1177  
PCT-US96-08140-1

## Query Match

Best Local Similarity 9.6%; Score 163.8; DB 5; Length 2075;

Matches 392; Conservative 0; Mismatches 332; Indels 9; Gaps 2;

QY 242 CGGTGACATGAGGGGAGCCACATGATCATCTTCTGGCCACTTCTTCTCTGCTCTTCT 301  
Db 135 CGCGCCAGTATGGGCGCGCTCGCGTCTCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 194  
QY 302 CTCATATGTTGTGTGCCAGCTGTGCGGACACCTGTGTCTCTCTCTCTCTCTCTCTCTCT 361  
Db 195 GCGGCGCTGTGCGGAGTGTGCGGCGCGCTGTGCGGCGCGCTGTGCGGCGCGCTGTGCGGCG 254  
QY 362 GTGCCACAGAGGGGTACCCCTGTGCTGTGATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 421  
Db 255 CTGCGCGCGCGCGGTGAGCTGTGCTGTGAGCGCTGTGCTGTGCGGCTGTGCGGCGCTGTGCGG 314  
QY 422 GAGGCTGGGGAGTCTCTGCGACACCTGTGCTGCGACCCCGACCCAGCCAGCGGCTGTGTTG 481  
Db 315 GCAGCTGGCGAGCTGTGACCGAGCGCGACCCCTGCGACCCCGACCAAGGCTCTCTCTCT 374  
QY 482 TCAGCTGGGGAGGCGCTGTGCGGCGCATGCGGCTGTGCTCTCTCTCTCTCTCTCTCTCTCT 541  
Db 375 TGACTTGGCTCTGCGGCGCAACGATCGCGCTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 431  
QY 542 TAGCTGTGAGGTGAATGGCGCGAGGTACCTGTGATGAGAGAGCTTTTAAACCCAAATTCGAG 601  
Db 432 TCCCTGCTACTTCGGTGTGATGACGGTGTGACCGCGAGGAGTCTCTTCCAGAGCAGCTGCAA 491  
QY 602 GGTCTGTGCGCTGTGATGACGGTGTGCTTACCTGCTGCGCTGTGCGCTGTGCGCTGTGCGATGT 661  
Db 492 GTACCAAGTCCAGTGTGATGACGGTGTGCGCTGTGCGCTGTGCGCTGTGCGCTGTGCGATGT 551  
QY 662 GCGGCTGCCAGTGTGAGTGTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 721  
Db 552 TCGTGTGCCAGCTGTGATGACGGTGTGCGCTGTGCGCTGTGCGCTGTGCGCTGTGCGCTGTGCG 611  
QY 722 CCCCAGTGGGTATGTGACCAAGGAGTGTGACCGGGGATCCAGGCTCCACGGCGCAAGG 781  
Db 612 CGAGGAGTGGGTGTGTGACGAGCCCAAGGACCAACCGTGTGGGCTTGCCTCGCGGC 671  
QY 782 ACACCAACTTCTGCCCTGTGATGACGGTGTGCGCTGTGAGTGTGAGTGTGAGTGTGAGTGT 835  
Db 672 TTACCGACTGGAGACAGCTTTGGCCGACGACCAACTATGATAGAGCCAACTGCCGTGT 731  
QY 836 TTGGAGCAGAGCTGTGATGACGGTGTGCGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 895  
Db 732 CCAGACCAAGAGTGTGATGACGGTGTGCGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 791  
QY 896 GTCCACCAAGAGTGTGATGACGGTGTGCGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 955  
Db 792 TACCAATGACAGCTGTGATGACGGTGTGCGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 851  
QY 956 CTGCTGCGAGCC 968  
Db 852 TTGCGAAGCTGAC 864

Search completed: July 28, 2003, 21:35:50  
Job time : 97.1244 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 15:36:41 ; Search time 421.038 seconds  
(without alignments)  
8368.853 Million cell updates/sec

Title: US-10-010-408-1  
Perfect score: 1708  
Sequence: 1 GACGCTTCTGATCTCCAGAG.....GCCTAGATAAACACCCAAA 1708

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1439767 seqs, 1031500376 residues  
Total number of hits satisfying chosen parameters: 2879534

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq1:\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq2:\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq3:\*
- 13: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query | Length | ID | Description        |
|------------|-------|-------|--------|----|--------------------|
| 1          | 1708  | 100.0 | 1708   | 15 | US-10-010-408-1    |
| 2          | 1278  | 74.8  | 1734   | 15 | US-10-112-267-17   |
| 3          | 1278  | 74.8  | 1734   | 15 | US-10-112-267-18   |
| 4          | 753   | 44.1  | 753    | 15 | US-10-010-408-3    |
| 5          | 681   | 39.9  | 681    | 15 | US-10-010-408-12   |
| 6          | 568.8 | 33.2  | 1337   | 11 | US-09-915-582-30   |
| 7          | 566.6 | 33.2  | 1352   | 11 | US-09-915-582-14   |
| 8          | 561.4 | 32.9  | 1266   | 14 | US-10-137-866-319  |
| 9          | 561.4 | 32.9  | 1266   | 14 | US-10-146-726-319  |
| 10         | 561.4 | 32.9  | 1266   | 14 | US-10-146-727-319  |
| 11         | 561.4 | 32.9  | 1266   | 14 | US-10-146-788-319  |
| 12         | 561.4 | 32.9  | 1266   | 14 | US-10-152-380-319  |
| 13         | 561.4 | 32.9  | 1266   | 14 | US-10-153-934-319  |
| 14         | 561.4 | 32.9  | 1266   | 15 | US-10-028-072-319  |
| 15         | 561.4 | 32.9  | 1266   | 15 | US-10-121-049-319  |
| 16         | 561.4 | 32.9  | 1266   | 15 | US-10-123-904-319  |
| 17         | 561.4 | 32.9  | 1266   | 15 | US-10-127-825A-319 |

|    |       |      |      |    |                    |                   |
|----|-------|------|------|----|--------------------|-------------------|
| 17 | 561.4 | 32.9 | 1266 | 15 | US-10-140-470-319  | Sequence 319, App |
| 18 | 561.4 | 32.9 | 1266 | 15 | US-10-175-746-319  | Sequence 319, App |
| 19 | 561.4 | 32.9 | 1266 | 15 | US-10-176-918-319  | Sequence 319, App |
| 20 | 561.4 | 32.9 | 1266 | 15 | US-10-176-921-319  | Sequence 319, App |
| 21 | 561.4 | 32.9 | 1266 | 15 | US-10-137-863-319  | Sequence 319, App |
| 22 | 561.4 | 32.9 | 1266 | 15 | US-10-140-474-319  | Sequence 319, App |
| 23 | 561.4 | 32.9 | 1266 | 15 | US-10-142-431-319  | Sequence 319, App |
| 24 | 561.4 | 32.9 | 1266 | 15 | US-10-143-114-319  | Sequence 319, App |
| 25 | 561.4 | 32.9 | 1266 | 15 | US-10-140-002-319  | Sequence 319, App |
| 26 | 561.4 | 32.9 | 1266 | 15 | US-10-142-419-319  | Sequence 319, App |
| 27 | 561.4 | 32.9 | 1266 | 15 | US-10-123-262-319  | Sequence 319, App |
| 28 | 561.4 | 32.9 | 1266 | 15 | US-10-142-423-319  | Sequence 319, App |
| 29 | 561.4 | 32.9 | 1266 | 15 | US-10-121-050-319  | Sequence 319, App |
| 30 | 561.4 | 32.9 | 1266 | 15 | US-10-141-755-319  | Sequence 319, App |
| 31 | 561.4 | 32.9 | 1266 | 15 | US-10-143-032-319  | Sequence 319, App |
| 32 | 561.4 | 32.9 | 1266 | 15 | US-10-123-108-319  | Sequence 319, App |
| 33 | 561.4 | 32.9 | 1266 | 15 | US-10-123-236-319  | Sequence 319, App |
| 34 | 561.4 | 32.9 | 1266 | 15 | US-10-123-261-319  | Sequence 319, App |
| 35 | 561.4 | 32.9 | 1266 | 15 | US-10-140-921-319  | Sequence 319, App |
| 36 | 561.4 | 32.9 | 1266 | 15 | US-10-140-928-319  | Sequence 319, App |
| 37 | 561.4 | 32.9 | 1266 | 15 | US-10-121-045-319  | Sequence 319, App |
| 38 | 561.4 | 32.9 | 1266 | 15 | US-10-123-292-319  | Sequence 319, App |
| 39 | 561.4 | 32.9 | 1266 | 15 | US-10-123-903-319  | Sequence 319, App |
| 40 | 561.4 | 32.9 | 1266 | 15 | US-10-124-819-319  | Sequence 319, App |
| 41 | 561.4 | 32.9 | 1266 | 15 | US-10-124-822-319  | Sequence 319, App |
| 42 | 561.4 | 32.9 | 1266 | 15 | US-10-140-925-319  | Sequence 319, App |
| 43 | 561.4 | 32.9 | 1266 | 15 | US-10-160-498-319  | Sequence 319, App |
| 44 | 561.4 | 32.9 | 1266 | 15 | US-10-124-824-319  | Sequence 319, App |
| 45 | 561.4 | 32.9 | 1266 | 15 | US-10-127-825A-319 | Sequence 319, App |

## ALIGNMENTS

## RESULT 1

US-10-010-408-1  
; Sequence 1, Application US/10010408  
; Publication No. US20020165185A1  
; GENERAL INFORMATION:

APPLICANT: John J. Castellet, Jr.  
TITLE OF INVENTION: No. US20020165185A1e1 Heparin-Induced CCN-Like Molecules and Uses Therefor

NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD, LLP  
STREET: 28 State Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/010.408  
FILING DATE: 07-Dec-2001  
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/044,273  
FILING DATE: March 19, 1998  
APPLICATION NUMBER: <Unknown>

ATTORNEY/AGENT INFORMATION:  
NAME: Amy E. Mandragouras

REGISTRATION NUMBER: 36,207  
REFERENCE/DOCKET NUMBER: MBI-004

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)742-4214

INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:

; LENGTH: 1708 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 249..1001  
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-10-010-408-1

Query Match 100.0%; Score 1708; DB 15; Length 1708;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1708; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGCTTCATGATCCAGAGGACCTGGGGTGGGACAGGGGCTTGGCAAGGCTCGAGCC 60  
DB 1 GACGCTTCATGATCCAGAGGACCTGGGGTGGGACAGGGGCTTGGCAAGGCTCGAGCC 60

QY 61 GCTGGGCAAGTGGCTTGGAAATGAGGCTCTTTATTAATGGAAGTGGAGGCTAAGAGGCTC 120  
DB 61 GCTGGGCAAGTGGCTTGGAAATGAGGCTCTTTATTAATGGAAGTGGAGGCTAAGAGGCTC 120

QY 121 CTGTGAGCTTGTCTTAAAGTCTTAGACACTTGTGGTGGCTTGGGCTTCACACACTGTGAGA 180  
DB 121 CTGTGAGCTTGTCTTAAAGTCTTAGACACTTGTGGTGGCTTGGGCTTCACACACTGTGAGA 180

QY 181 CACCTTCGTGGTGGGCTCCAGGCTCCAGGCTCACCTTCAGGTTTGAAGCTGGCTTCCACAAAGGAC 240  
DB 181 CACCTTCGTGGTGGGCTCCAGGCTCCAGGCTCACCTTCAGGTTTGAAGCTGGCTTCCACAAAGGAC 240

QY 241 ACGGTGACATGAGGGGAGCCGACATGATCCATCTTCTGGCCACTTCTTCTCTCTCCCTTC 300  
DB 241 ACGGTGACATGAGGGGAGCCGACATGATCCATCTTCTGGCCACTTCTTCTCTCTCCCTTC 300

QY 301 TCTCAATGGTGTGCTCCAGCTGTGCGGACACCTGTACCTGTCTTGGACACACCCG 360  
DB 301 TCTCAATGGTGTGCTCCAGCTGTGCGGACACCTGTACCTGTCTTGGACACACCCG 360

QY 361 AGTGGCCACAGGGGTACCCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 420  
DB 361 AGTGGCCACAGGGGTACCCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 420

QY 421 GGAGGCTGGGAGTCTGCGACACCTGATGATGCTGCGACCCGACGAGGCGCTGGTTT 480  
DB 421 GGAGGCTGGGAGTCTGCGACACCTGATGATGCTGCGACCCGACGAGGCGCTGGTTT 480

QY 481 GTCAGCTGGGACGCGCTGCGGCGCATGGGCTGTGTCTCTTGGATGAGGATGAGC 540  
DB 481 GTCAGCTGGGACGCGCTGCGGCGCATGGGCTGTGTCTCTTGGATGAGGATGAGC 540

QY 541 GTAGCTGTGAGTGAATGGCCGACGATACCTGGATGGAGAGACCTTTAAACCCAAATTGCA 600  
DB 541 GTAGCTGTGAGTGAATGGCCGACGATACCTGGATGGAGAGACCTTTAAACCCAAATTGCA 600

QY 601 GGGTCTGTGCGCTGTGATGACGCTGGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTG 660  
DB 601 GGGTCTGTGCGCTGTGATGACGCTGGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTG 660

QY 661 TCGGCTGCCAGCTGGGACTGCCACGCGCCCAAGAAATACAGGTGCCAGGAAAGTCT 720  
DB 661 TCGGCTGCCAGCTGGGACTGCCACGCGCCCAAGAAATACAGGTGCCAGGAAAGTCT 720

QY 721 GCCCGAGTGGTATGTGACCAAGGAGTGACACCGGCGATCCAGCGCTCCAGCGGCGCAAG 780  
DB 721 GCCCGAGTGGTATGTGACCAAGGAGTGACACCGGCGATCCAGCGCTCCAGCGGCGCAAG 780

QY 781 GACACCAACTTCTGCCCTGTGCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840  
DB 781 GACACCAACTTCTGCCCTGTGCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840

QY 841 GCACAGCTGGGGGCGCTGCTCAACCACTGTGGGCTGGGCTAGCCACCGGAGTGTCCA 900  
DB 841 GCACAGCTGGGGGCGCTGCTCAACCACTGTGGGCTGGGCTAGCCACCGGAGTGTCCA 900

## RESULT 2

US-10-112-267-17

; Sequence 17, Application US/10112267

; Publication No. US2003006878A1

; GENERAL INFORMATION:

; APPLICANT: Botstein, David A.

; APPLICANT: Cohen, Robert

; APPLICANT: Goddard, Audrey

; APPLICANT: Gurney, Austin L.

; APPLICANT: Hillan, Kenneth J.

; APPLICANT: Lawrence, David A.

; APPLICANT: Levine, Arnold J.

; APPLICANT: Pennica, Diane

; APPLICANT: Roy, Margaret Ann

DB 841 GCACAGCTGGGGGCGCTGCTCAACCACTGTGGCTGGGCAATAGCCACCCGAGTGTCCA 900  
QY 901 ACCAGAACGATTTGCGCAACTGGAGATCCAAACGCGGCTGTGTGTCGCCAGACCTGCG 960  
DB 901 ACCAGAACGATTTGCGCAACTGGAGATCCAAACGCGGCTGTGTGTCGCCAGACCTGCG 960  
QY 961 TGGCAGCCAGGAGCCACAGCTCATGGAACAGTGTCTTCTAAGGCCAACTGGGGATGCGGA 1020  
DB 961 TGGCAGCCAGGAGCCACAGCTCATGGAACAGTGTCTTCTAAGGCCAACTGGGGATGCGGA 1020  
QY 1021 TACAGGGCTGCTCATCTCCAGCAATGACCCCTAGACAGGCGCTGGAGTGTGTAGAT 1080  
DB 1021 TACAGGGCTGCTCATCTCCAGCAATGACCCCTAGACAGGCGCTGGAGTGTGTAGAT 1080  
QY 1081 GCTCTTCTCCAGTCTTGGGCTGACGTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140  
DB 1081 GCTCTTCTCCAGTCTTGGGCTGACGTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140  
QY 1141 ACTGAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200  
DB 1141 ACTGAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200  
QY 1201 GTTCAGCCTGGAAATTTCTGGGTTCTCCTGGCTCATTTCTCCTCAAAACATCCCTGTACA 1260  
DB 1201 GTTCAGCCTGGAAATTTCTGGGTTCTCCTGGCTCATTTCTCCTCAAAACATCCCTGTACA 1260  
QY 1261 GACAAACAAAGACCTTTAAACCTAGGCTATGCTAGGCTGGGCAACCTGGGCAACCTGGG 1320  
DB 1261 GACAAACAAAGACCTTTAAACCTAGGCTATGCTAGGCTGGGCAACCTGGGCAACCTGGG 1320  
QY 1321 GATAAGTCAATGTTAGGACACGACAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380  
DB 1321 GATAAGTCAATGTTAGGACACGACAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380  
QY 1381 ACTTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440  
DB 1381 ACTTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440  
QY 1441 TGAGAACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500  
DB 1441 TGAGAACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500  
QY 1501 GATGAATCACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560  
DB 1501 GATGAATCACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560  
QY 1561 CAGGAATCACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620  
DB 1561 CAGGAATCACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620  
QY 1621 AGGGCTTTCTTCTCAGCATGAGAAAGACAAAGGGAGGCTCCACACCTCTCTCTCTGAGGA 1680  
DB 1621 AGGGCTTTCTTCTCAGCATGAGAAAGACAAAGGGAGGCTCCACACCTCTCTCTCTGAGGA 1680  
QY 1681 CTAGTCTAGCTAGTAATAACACCCAAA 1708  
DB 1681 CTAGTCTAGCTAGTAATAACACCCAAA 1708

## RESULT 2

US-10-112-267-17

; Sequence 17, Application US/10112267

; Publication No. US2003006878A1

; GENERAL INFORMATION:

; APPLICANT: Botstein, David A.

; APPLICANT: Cohen, Robert

; APPLICANT: Goddard, Audrey

; APPLICANT: Gurney, Austin L.

; APPLICANT: Hillan, Kenneth J.

; APPLICANT: Lawrence, David A.

; APPLICANT: Levine, Arnold J.

; APPLICANT: Pennica, Diane

; APPLICANT: Roy, Margaret Ann

APPLICANT: Wood, William I.  
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: P1176R2  
; CURRENT APPLICATION NUMBER: US/10/112,267  
; CURRENT FILING DATE: 2002-03-27  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/182,145B  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-29  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/063,704  
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-10-29  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/073,612  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-04  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/081,695  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-14  
; NUMBER OF SEQ ID NOS: 156  
; SEQ ID NO 17  
; LENGTH: 1734  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-10-112-267-17

Query Match 74.8%; Score 1278; DB 15; Length 1734;  
Best Local Similarity 88.5%; Pred. No. 0;  
Matches 1520; Conservative 0; Mismatches 165; Indels 33; Gaps 11;

Qy 3 CGCTTCTGATCTCCAGAGGACCCCTGGGGTGGGACAGGGGCTTGGCAAGGCTGCGAGCCG 62  
Db 13 CGCTCTGATCTCCAGAGGACCCCGGGCTGGGACAGGGGCTTGGCGAGGCTGCGAGCTG 72

Qy 63 TG-GGCAGTGGCTTGGGAATGAGTCTTTACTGGGAACCTGAGGAGCTTAAGAGGCTCC 121  
Db 73 TGTGGCAGTGGCTTGGGAATGAGTCTTTCTTGGTGGGAACCTGAGGAGCTGAGAGGCTCC 132

Qy 122 TGTGAG---CTTGTCTTAAAGTCTTAGCACTTTGTGGTGGCTTGGGCTTTCACACACTGCA 178  
Db 133 TGTGAGGCTCTGCTCTTAAAGTCTTAGCACTTTGTGGTGGCTTGGGCTTTCACACACTGCA 192

Qy 179 GACACCTTGTGGTGGCTCCACGGCTTCACTTCACTTGAAGCTTGAAGCTTGCACAAAGG 238  
Db 193 GACACCTTGTGGTGGCTTCTCGGCC-----TCAGTTTGAAGCTTGGCTTGCACAAAGG 246

Qy 239 ACAGGTGACATGAGGGGACCCACCTGATCCATCTTCTGCGCACTTCTTCTCTGCT 298  
Db 247 ACAGGTGACATGAGGGGACCCACCTGATCCATCTTCTGCGCACTTCTTCTCTGCT 306

Qy 299 TCTCTCAATGCTGTGCCAGCTGTGCCGACACCCCTGTACCTGTCTTGGACACCAAC 358  
Db 307 TCTCTCAATGCTGTATCCAGCTGTGCCGACACCCCTGTGCTTGGACACCAAC 366

Qy 359 CCAGTGGCCACAGGGGTACCCCTGTGCTGGATGGCTGTGGCTGTGCTAAAGTGTGC 418  
Db 367 CCAGTGGCCACAGGGGTACCCCTGTGCTGGATGGCTGTGGCTGTGCTAAAGTGTGC 426

Qy 419 ACAGAGCTGGGGAGTCTGCGACACCTGCTGCTGCGACCCAGCCAGCCAGGCTGTGT 478  
Db 427 ACAGAGCTGGGGAGTCTGCGACACCTGCTGCTGCGACCCAGCCAGGCTGTGT 486

Qy 479 TTGTGACCTGGGGAGGCTTGGCGGCTGTGTGCTTCTTGGATGAGGATGA 538  
Db 487 TTGTGACCTGGGGAGGCTTGGCGGCTGTGTGCTTCTTGGATGAGGATGA 546

Qy 539 CGGTAGCTGTGAGTGAATGGCGCAGGTACCTGGATGGAGAGACCTTTAAACCAATTG 598  
Db 547 CGGGAGCTGTGAGTGAATGGCGCAGGTACCTGGATGGGAGACCTTTAAACCAATTG 606

Qy 599 CAGGCTCTGCTGCTGATGACGCTGCTTCACTGCTGCCGCTGTGCGATGAGGA 658  
Db 607 CAGGCTCTGCTGCTGATGACGCTGCTTCACTGCTGCCGCTGTGCGATGAGGA 666

Qy 659 TGTGCGGCTGCCAGCTGGGACTGCCACCGCCCAAGAAATACAGGTGCCAGGAAAGT 718  
Db 667 TGTGCGGCTGCCAGCTGGGACTGCCACCGCCCAAGAAATACAGGTGCCAGGAAAGT 726

Qy 719 CTGCCCCAGTGGGTATGTGACCAAGGAGTGA---CACCGGCGATCCAGCGCTCCAGCGC 775

RESULT 3

US-10-112-267-18/c

; Sequence 18, Application US/10112267

; Publication No. US20030068678A1

; GENERAL INFORMATION:

Db 727 CTGCCCGAGTGGGTGTGACCAAGCAGTGTGACCGCGCAATCCAGCCCTCTCCAGC 786  
Qy 776 GCAAGGACACAACTTTCTGCCCTTGTCACTCTCTGCTGTGCTGTGCTCTCTTGTCTCAA 835  
Db 787 CCAAGGACACAACTTTCTGCCCTTGTCACTCTCTGCTGTGCTGTGCTGTGCTCTCAA 846  
Qy 836 TTGGAGCACAGCCTGGGGCCCTGCTCAACACCTGTGGGCTGGGATAGCCACCCAGT 895  
Db 847 CTGGAGCACAGCCTGGGGCCCTGCTCAACACCTGTGGGTTGGGATAGCCACCCAGT 906  
Qy 896 GTCCAAACAGAACCGATTCTGCAACTGGGATCCAAACCGCTGTGCTGTGCTGCTGCTG 955  
Db 907 ATCCAAACAGAACCGATTCTGCAACTGGGATCCAAACCGCTGTGCTGTGCTGCTGCTG 966  
Qy 956 CTGCTGGCAGCCAGGAGCCACAGCTCATGGAACAGTGTCTTAAGGCCA-ACTGGGGA 1014  
Db 967 CTGCTGGCAGCCAGGAGCCACAGCTCATGGAACAGTGTCTTAAGGCCA-ACTGGGGA 1026  
Qy 1015 TGGGATACAGGGCCTGCCATCTCAGCAATGACCTAGGACAGGCGCTGGAGTCTG 1074  
Db 1027 TGTGATACAGGGCCTGCCATCTCAGCAATGACCTAGGACAGGCGCTGGAGTCTG 1086  
Qy 1075 GTAGATGCTCTTCTCATGCTCTTGGCTGCAAGTTAACTGTCTTGGTGTGCTGTGCTGT 1134  
Db 1087 GTAGATGCTCTTCTCATGCTCTTGGCTGCAAGTTAACTGTCTTGGTGTGCTGTGCTGT 1146  
Qy 1135 AGAGCCACTGAGGAGTCCCTGCTGCTGAGGTAGGAGGAGGAGGAGGAGGAGGAGGAGG 1194  
Db 1147 AGAGCCCTGAGGAGTCCCTGCTGCTGAGGTAGGAGGAGGAGGAGGAGGAGGAGGAGG 1206  
Qy 1195 TCTCTG--TTCAGCCCTGGAAATCTGGGTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1252  
Db 1207 TCTCTGATCTGACCCAGGCTTCTGGGTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1266  
Qy 1253 ACAAAGGAGACACCAAAAGACCTTTAAACCTAGGCTATGCTGCTGCTGCTGCTGCTGCTG 1312  
Db 1267 ATGAAAGGAGACCAAAAGGAGCTTTAAAGCTTAAGCTGTGCTGCTGCTGCTGCTGCTG 1326  
Qy 1313 GTCTGGGGTAAGTCAATGTTAGG-ACCAGACAGCAGATTCCTGAAACTTCCAAATTC 1371  
Db 1327 ATGCTGGGATAGTACAGTAATAGTACCAGCAGCAGATTCCTGAAACTTCCAAATTC 1386  
Qy 1372 CCTTCTGGACTTCTGTATGCTTGTCCCAAGATGATGAATGAATGCTGTAAGTGTACCT 1431  
Db 1387 CCTTCTGGACTTCTGTATGCTTGT-CCCAAGATTAAGTGTGCTGCTGCTGCTGCTGCT 1445  
Qy 1432 TCCCTGACCTGAGAACCCCTGCTGCTGCGGAAGTATTCAGGGGAGAAATTCCTGTGA 1491  
Db 1446 TTCTGATCTGAGAACCCCTGCTGCTGCGGCTGGG-AGAAATTTCTGGGA 1492  
Qy 1492 ACATGAAGAGAT-GAATCACTGCTCTTAAGAAATTCCTGAAAGTCCAGAACTTTCAGC 1550  
Db 1493 ACATGAAGAGATGAATCACTATTCCTTAAGAGCTTTGCCAAGTCCAGAACTTTCAGC 1552  
Qy 1551 TTTGTATTTTTCAGGAATGACATCTCTTAAGACTCTGCAAAACAGGAGGCTCCACACT 1610  
Db 1553 TTTGTATTTTGAATAATACATCTCTTAAATGCTCACAAGCAAG-AGGCTCCACACTT 1611  
Qy 1611 CTGGCAGGCGAGGCTTTCTTTCAGCATGAGAAAGAGGAGGAGGAGGAGGAGGAGGAGG 1670  
Db 1612 CTGGCAGGCGAGGCTTTCTTTCAGCATGAGAGAGGAGGAGGAGGAGGAGGAGGAGG 1671  
Qy 1671 CTCTGGAGGAGTGTAGCTTGTAGCTAGAAATAAACCCCAAA 1708  
Db 1672 CTCTGGAGGAGTGGCCGGTCTGGAATAAACCCCAAA 1709

APPLICANT: Botstein, David A.  
APPLICANT: Cohen, Robert  
APPLICANT: Goddard, Audrey  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: Lawrence, David A.  
APPLICANT: Levine, Arnold J.  
APPLICANT: Pennica, Diane  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME  
FILE REFERENCE: P1176R2  
CURRENT APPLICATION NUMBER: US/10/112,267  
PRIOR FILING DATE: 2002-03-27  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/182,145B  
PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-29  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/063,704  
PRIOR FILING DATE: EARLIER FILING DATE: 1997-10-29  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/073,612  
PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-04  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/081,695  
PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-14  
NUMBER OF SEQ ID NOS: 156  
SEQ ID NO 18  
LENGTH: 1734  
TYPE: DNA  
ORGANISM: Mus musculus  
US-10-112-267-18

Query Match 74.8%; Score 1278; DB 15; Length 1734;  
Best Local Similarity 88.5%; Pred. No. 0;  
Matches 1520; Conservative 0; Mismatches 165; Indels 33; Gaps 11;

QY 3 CGCTTCTGATCCAGAGACCTCGGTGGGACAGGGCCCTTGCAAGGCTGCAGCGC 62  
DB 1722 CGCTCCTGATCTCCAGAGACCCCGGCTGGGACAGGGCCCTTGCGAGGCTGCAGCTGC 1663

QY 63 TGGCAGTGGCTTGGATGGAGGCTTTTATTACTGGGAACCTAGAGGCTCC 121  
DB 1662 TGTGGCAGTAGCTTGGGATGGAGGCTTTCTTGTCTGGGAACCTAGAGGCTCC 1603

QY 122 TGTGAG--CTTGTCTTAAGCTTAGCAGTGTGTGGTGGCTTGGCTTGCACACCTGCA 178  
DB 1602 TGTGAGCTCTGTCTTAAGCTTAGCAGTGTGTGGTGGCTTGGCTTGCACACCTGCA 1543

QY 179 GACACCTTGTGGTGGCTTCCAGGCTCCACCTTTCAGGTTTGAAGCTGGCTCCACAGGG 238  
DB 1542 GACACCTTGTGGTGGCTTCCAGGCTCCACCTTTCAGGTTTGAAGCTGGCTCCACAGGG 1489

QY 239 ACACGGTGACATGAGGGGAGCCCACTGATCCATCTTCTGGCCACTTCTCTCTGCTGCT 298  
DB 1488 ACACGGTGACATGAGGGGAGCCCACTGATCCATCTTCTGGCCACTTCTCTCTGCTGCT 1429

QY 299 TCTCTCAATGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 358  
DB 1428 TCTCTCAATGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1369

QY 359 CCAGTGGCCACAGGGGGTACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 418  
DB 1368 CCAGTGGCCACAGGGGGTACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1309

QY 419 ACGGAGGCTGGGGAGTCTCTGGGACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 478  
DB 1308 ACGGAGGCTGGGGAGTCTCTGGGACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1249

QY 479 TTGTGAGCTGGGGAGGCTTGGGCGGCGCATGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 538  
DB 1248 TTGTGAGCTGGGGAGGCTTGGGCGGCGCATGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1189

QY 539 CGGTAGCTGTGAGGTGAATGGCGCGAGGTACCTGATGAGAGACCTTTAAACCCAAATTG 598  
DB 1188 CGGAGCTGTGAGGTGAATGGCGCGAGGTACCTGATGAGAGACCTTTAAACCCAAATTG 1129

QY 599 CAGGCTCTGTGCGCCTGTGATGACGGTGGCTTACCTGCTGCGCTGTGCAAGTGA 658  
DB 1128 CAGGCTTTTGTGCGCCTGTGATGACGGTGGCTTACCTGCTGCGCTGTGCAAGTGA 1069

QY 659 TGTGCGGCTGCGCAGCTGGGACTGCCACGCCCCCAAGAGAATACAGGTGCCAGGAAGTG 718  
DB 1068 TGTGCGGCTGCGCAGCTGGGACTGCCACGCCCCCAAGAGAATACAGGTGCCAGGAAGTG 1009

QY 719 CTGCCCCAGTGGGTATGTGACCCAGGAGTGA--CACCAGGCTATCCAGCGCTCCACGCG 775  
DB 1008 CTGCCCCAGTGGGTATGTGACCCAGGAGTGA--CACCAGGCTATCCAGCGCTCCACGCG 949

QY 776 GCAAGGACCAACATTTCTGCCCCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 835  
DB 948 CCAAGGACCAACATTTCTGCCCCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 889

QY 836 TTGGAGCACAGCTGGGGCCCTGCTCAACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 895  
DB 888 CTGGAGCACAGCTGGGGCCCTGCTCAACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 829

QY 896 GTCCAAACCAACATTTCTGCCCCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 955  
DB 828 ATCCAAACCAACATTTCTGCCCCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 769

QY 956 CTGCTGCGAGCGGAGGACGCTCATGGAACAGTCTTCTTAAGGCCA--ACTGGGGA 1014  
DB 768 CTGCTGCGATCCAGGAGGACGCTCATGGAACAGTCTTCTTAAGGCCA--ACTGGGGA 709

QY 1015 TGGGATACAGGCTGCGCATCTCCAGCAATGACCTTAGGACAGGCGCTGGAGCTGCTG 1074  
DB 708 TGTGGATACAGGCTGCGCATCTCCAGCAATGACCTTAGGACAGGCGCTGGAGCTGCTG 649

QY 1075 GTAGATGCTCTTCTCCATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1134  
DB 648 GTAGATGCTCTTCTCCATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 589

QY 1135 AGAGCCATGAGGAGTATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1194  
DB 588 AGAGCCATGAGGAGTATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 529

QY 1195 TCTCTGG--TTCAGGCTGGAATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1252  
DB 528 TCTCTGGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 469

QY 1253 ACAAAAGGACACCAAAAGACCTTTAAACCTAGGCTATATCTGCGGCAAACTGGCCACC 1312  
DB 468 ATGAAAGGACACCAAAAGACCTTTAAAGCTAGCTGCTGCTGCGCAAGCTGGCCACC 409

QY 1313 GTGCTGGGATGAGTCAATGTTAGG--ACCAGACAGCAGATTGCTGAACTTCCAATTC 1371  
DB 408 ATGCTGGGATGAGTCAATGTTAGG--ACCAGACAGCAGATTGCTGAACTTCCAATTC 349

QY 1372 CTTCTTGGACTTGTATGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1431  
DB 348 CTTCTTGGACTTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 290

QY 1432 TCCCTGACTGAGAACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1491  
DB 289 TCCCTGACTGAGAACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 243

QY 1492 ACATGAAGAGAT--GAACTGACTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1550  
DB 242 ACATGAAGAGATGAACTGACTGCTTCTTAAAGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 183

QY 1551 TTTGTATTTCAGGAAATGACATCTTTAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1610  
DB 182 TTTGTATTTCAGGAAATGACATCTTTAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 124

QY 1611 CTGGCAGGCGGCTTTCTTCTTCCAGTATGAGAAAGACAGGAGGACAGCAGACTACTCTC 1670  
DB 123 CTGGCAGGCGGCTTTCTTCTTCCAGTATGAGAGACAGGAGGACAGCAGACTACTCTC 64

QY 1671 CTCTGGAGCTAGTCTAGCCTAGAAATTAACACCCAAA 1708





TELEPHONE: (617)227-7400  
TELEFAX: (617)742-4214  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 681 base pairs  
TYPE: nucleic acid  
STRADEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..681  
SEQUENCE DESCRIPTION: SEQ ID NO: 12:  
US-10-010-408-12

Query Match 39.9%; Score 681; DB 15; Length 681;  
Best Local Similarity 100.0%; Pred. No. 1.1e-202;  
Matches 681; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 318 CAGCTGTGCGGACACCTGTACCTGTCTTGGACACACCCAGTGGCCACAGGGGTA 377  
Db |||||  
Db 1 CAGCTGTGCGGACACCTGTACCTGTCTTGGACACACCCAGTGGCCACAGGGGTA 60  
QY 378 CCCTGGTCTGGATGGCTGGCTGCTGCTAAAGTGTGACAGGAGGCTGGGGAGTCC 437  
Db |||||  
Db 61 CCCCTGGTCTGGATGGCTGGCTGCTGCTAAAGTGTGACAGGAGGCTGGGGAGTCC 120  
QY 438 TGGACACCTGCATGTCTGCGACCCAGCCAGGCGCTGTTGTGAGCTGGGCGAGC 497  
Db |||||  
Db 121 TGGACACCTGCATGTCTGCGACCCAGCCAGGCGCTGTTGTGAGCTGGGCGAGC 180  
QY 498 CTGGCGGCGCATGGGCTGTGTCTCTTGGATGAGGATGACGGTGTGAGGTGAAT 557  
Db |||||  
Db 181 CTGGCGGCGCATGGGCTGTGTCTCTTGGATGAGGATGACGGTGTGAGGTGAAT 240  
QY 558 GCGCGAGTACTGTGATGGAGAGACCTTTAAACCAATTCAGGCTGCTGCGCGTGT 617  
Db |||||  
Db 241 GCGCGAGTACTGTGATGGAGAGACCTTTAAACCAATTCAGGCTGCTGCGCGTGT 300  
QY 618 GATGACGCTGCTTACCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 677  
Db |||||  
Db 301 GATGACGCTGCTTACCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360  
QY 678 GACTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 737  
Db |||||  
Db 361 GACTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 420  
QY 738 GACGAGGAGTACACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 797  
Db |||||  
Db 421 GACGAGGAGTACACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480  
QY 798 CTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 857  
Db |||||  
Db 481 CTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540  
QY 858 TGTCAACACCTGTGGGCTGGGCTAGACACCCAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 917  
Db |||||  
Db 541 TGTCAACACCTGTGGGCTGGGCTAGACACCCAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 600  
QY 918 CAACCTGGAGATCAACCGGCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 977  
Db |||||  
Db 601 CAACCTGGAGATCAACCGGCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660  
QY 978 AGCTCATGGAACAGTGTTC 998  
Db |||||  
Db 661 AGCTCATGGAACAGTGTTC 681

## RESULT 6

US-09-915-582-30  
Sequence 30, Application US/09915582  
Patent No. US20020120103A1  
GENERAL INFORMATION:

APPLICANT: Rosen et al.  
TITLE OF INVENTION: 17 Human Secreted Proteins  
FILE REFERENCE: PS723P1  
CURRENT APPLICATION NUMBER: US/09/915,582  
CURRENT FILING DATE: 2001-07-27  
PRIOR APPLICATION NUMBER: PCT/US01/01431  
PRIOR FILING DATE: 2001-01-17  
PRIOR APPLICATION NUMBER: 60/179,065  
PRIOR FILING DATE: 2000-01-31  
PRIOR APPLICATION NUMBER: 60/180,628  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: 60/231,968  
PRIOR FILING DATE: 2000-09-12  
NUMBER OF SEQ ID NOS: 97  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 30  
LENGTH: 1337  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (1337)  
OTHER INFORMATION: n equals a,t,g, or c  
US-09-915-582-30

Query Match 33.2%; Score 566.8; DB 11; Length 1337;  
Best Local Similarity 77.0%; Pred. No. 8.4e-167;  
Matches 728; Conservative 3; Mismatches 205; Indels 10; Gaps 3;

QY 196 CTCACAGGCTCACCTTCAGGTTTGAAGTGTGCTCCACAAAGGACACCGTGCATGAGGG 255  
Db |||||  
Db 2 CTCACAGGCTTCAGGTTTGAAGTGTGCTCCACAAAGGACACCGTGCATGAGGG 53  
QY 256 GCAGCCCACTGATCATCTTCGCGCACTTCCTCTCTGCTCTCTCTCTCTCTCTCTCTCT 315  
Db |||||  
Db 54 GCACACCAAGACCCACCTCTCTGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 113  
QY 316 CCCAGCTGTGCGGACACCTCTTACCTGTCTTGGACACCAACCCAGTGCACAGGGGG 375  
Db |||||  
Db 114 CCCAGCTGTGCGGACACCACTTACCTGTCTCTGCGCACTTCCTGCGGCTGCGGAG 173  
QY 376 TACCCCTGGTGTGATGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 435  
Db |||||  
Db 174 TACCCCTGGTGTGATGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 233  
QY 436 COTGCGACACCTGCATGTCTGCGACCCAGCAGCGGCTGTGCTGTGCTGTGCTGTGCTGTGCT 495  
Db |||||  
Db 234 COTGCGACACCTGCATGTCTGCGACCCAGCAGCGGCTGTGCTGTGCTGTGCTGTGCTGTGCT 293  
QY 496 GOCCTGGCGGCAATGGGCTGTGCTGTCTTGGATGAGGATGACGGTGTGCTGTGCTGTGCTGTGCT 555  
Db |||||  
Db 294 GACCCGGTGMCGGGGGGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 353  
QY 556 ATGGCGCGGAGTACCTGTGATGGAGAGACCTTTAAACCAATTCAGAGGTCCTGTGCTGCTGTGCT 615  
Db |||||  
Db 354 AGCGCGGCTGTATCGGAAGGGAGACCTTTCCAGCCCACTGTCAGACATCCGCTGCTGCTGCTGCT 413  
QY 616 GTGATGACGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 675  
Db |||||  
Db 414 GCGAGGAGCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 473  
QY 676 GGGACTGCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 735  
Db |||||  
Db 474 GGGACTGCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 533  
QY 736 GTGACAGGAGTGTGACACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 795  
Db |||||  
Db 534 GCGGCAAGGAGGGGAGGAGTGGGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 593  
QY 796 CCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 855  
Db |||||  
Db 594 GCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 653



1 PRIOR FILING DATE: 1997-08-26  
2 PRIOR APPLICATION NUMBER: 60/059113  
3 PRIOR FILING DATE: 1997-09-17  
4 PRIOR APPLICATION NUMBER: 60/059115  
5 PRIOR FILING DATE: 1997-09-17  
6 PRIOR APPLICATION NUMBER: 60/059117  
7 PRIOR FILING DATE: 1997-09-17  
8 PRIOR APPLICATION NUMBER: 60/059122  
9 PRIOR FILING DATE: 1997-09-17  
10 PRIOR APPLICATION NUMBER: 60/059184  
11 PRIOR FILING DATE: 1997-09-17  
12 PRIOR APPLICATION NUMBER: 60/059263  
13 PRIOR FILING DATE: 1997-09-18  
14 PRIOR APPLICATION NUMBER: 60/059352  
15 PRIOR FILING DATE: 1997-09-19  
16 PRIOR APPLICATION NUMBER: 60/059588  
17 PRIOR FILING DATE: 1997-09-19  
18 PRIOR APPLICATION NUMBER: 60/059836  
19 PRIOR FILING DATE: 1997-09-24  
20 PRIOR APPLICATION NUMBER: 60/062250  
21 PRIOR FILING DATE: 1997-10-17  
22 PRIOR APPLICATION NUMBER: 60/062285  
23 PRIOR FILING DATE: 1997-10-17  
24 PRIOR APPLICATION NUMBER: 60/062287  
25 PRIOR FILING DATE: 1997-10-17  
26 PRIOR APPLICATION NUMBER: 60/062814  
27 PRIOR FILING DATE: 1997-10-24  
28 PRIOR APPLICATION NUMBER: 60/062816  
29 PRIOR FILING DATE: 1997-10-24  
30 PRIOR APPLICATION NUMBER: 60/063045  
31 PRIOR FILING DATE: 1997-10-24  
32 PRIOR APPLICATION NUMBER: 60/063082  
33 PRIOR FILING DATE: 1997-10-31  
34 PRIOR APPLICATION NUMBER: 60/063127  
35 PRIOR FILING DATE: 1997-10-24  
36 PRIOR APPLICATION NUMBER: 60/063327  
37 PRIOR FILING DATE: 1997-10-27  
38 PRIOR APPLICATION NUMBER: 60/063329  
39 PRIOR FILING DATE: 1997-10-27  
40 PRIOR APPLICATION NUMBER: 60/063550  
41 PRIOR FILING DATE: 1997-10-28  
42 PRIOR APPLICATION NUMBER: 60/063561  
43 PRIOR FILING DATE: 1997-10-28  
44 PRIOR APPLICATION NUMBER: 60/063704  
45 PRIOR FILING DATE: 1997-10-29  
46 PRIOR APPLICATION NUMBER: 60/063733  
47 PRIOR FILING DATE: 1997-10-29  
48 PRIOR APPLICATION NUMBER: 60/063735  
49 PRIOR FILING DATE: 1997-10-29  
50 PRIOR APPLICATION NUMBER: 60/063738  
51 PRIOR FILING DATE: 1997-10-29  
52 PRIOR APPLICATION NUMBER: 60/063755  
53 PRIOR FILING DATE: 1997-10-17  
54 PRIOR APPLICATION NUMBER: 60/064248  
55 PRIOR FILING DATE: 1997-11-03  
56 PRIOR APPLICATION NUMBER: 60/064809  
57 PRIOR FILING DATE: 1997-11-07  
58 PRIOR APPLICATION NUMBER: 60/065186  
59 PRIOR FILING DATE: 1997-11-12  
60 PRIOR APPLICATION NUMBER: 60/065846  
61 PRIOR FILING DATE: 1997-11-17  
62 PRIOR APPLICATION NUMBER: 60/066364  
63 PRIOR FILING DATE: 1997-11-21  
64 PRIOR APPLICATION NUMBER: 60/066453  
65 PRIOR FILING DATE: 1997-11-24  
66 PRIOR APPLICATION NUMBER: 60/066511  
67 PRIOR FILING DATE: 1997-11-24  
68 PRIOR APPLICATION NUMBER: 60/066770  
69 PRIOR FILING DATE: 1997-11-24  
70 PRIOR APPLICATION NUMBER: 60/069212  
71 PRIOR FILING DATE: 1997-12-11  
72 PRIOR APPLICATION NUMBER: 60/069278  
73 PRIOR FILING DATE: 1997-12-11

74 PRIOR APPLICATION NUMBER: 60/069334  
75 PRIOR FILING DATE: 1997-12-11  
76 PRIOR APPLICATION NUMBER: 60/069694  
77 PRIOR FILING DATE: 1997-12-16  
78 PRIOR APPLICATION NUMBER: 60/072320  
79 PRIOR FILING DATE: 1998-01-23  
80 PRIOR APPLICATION NUMBER: 60/073612  
81 PRIOR FILING DATE: 1998-02-04  
82 PRIOR APPLICATION NUMBER: 60/074086  
83 PRIOR FILING DATE: 1998-02-09  
84 PRIOR APPLICATION NUMBER: 60/074092  
85 PRIOR FILING DATE: 1998-02-09  
86 PRIOR APPLICATION NUMBER: 60/077791  
87 PRIOR FILING DATE: 1998-03-12  
88 PRIOR APPLICATION NUMBER: 60/078910  
89 PRIOR FILING DATE: 1998-03-20  
90 PRIOR APPLICATION NUMBER: 60/079294  
91 PRIOR FILING DATE: 1998-03-25  
92 PRIOR APPLICATION NUMBER: 60/079663  
93 PRIOR FILING DATE: 1998-02-27  
94 PRIOR APPLICATION NUMBER: 60/079728  
95 PRIOR FILING DATE: 1998-03-27  
96 PRIOR APPLICATION NUMBER: 60/080165  
97 PRIOR FILING DATE: 1998-03-31  
98 PRIOR APPLICATION NUMBER: 60/081203  
99 PRIOR FILING DATE: 1998-04-09  
100 PRIOR APPLICATION NUMBER: 60/081229  
101 PRIOR FILING DATE: 1998-04-09  
102 PRIOR APPLICATION NUMBER: 60/081695  
103 PRIOR FILING DATE: 1998-04-14  
104 PRIOR APPLICATION NUMBER: 60/081817  
105 PRIOR FILING DATE: 1998-04-15  
106 PRIOR APPLICATION NUMBER: 60/081818  
107 PRIOR FILING DATE: 1998-04-15  
108 PRIOR APPLICATION NUMBER: 60/082999  
109 PRIOR FILING DATE: 1998-04-24  
110 PRIOR APPLICATION NUMBER: 60/083322  
111 PRIOR FILING DATE: 1998-04-28  
112 PRIOR APPLICATION NUMBER: 60/083545  
113 PRIOR FILING DATE: 1998-04-29  
114 PRIOR APPLICATION NUMBER: 60/084600  
115 PRIOR FILING DATE: 1998-05-07  
116 PRIOR APPLICATION NUMBER: 60/084627  
117 PRIOR FILING DATE: 1998-05-07  
118 PRIOR APPLICATION NUMBER: 60/084637  
119 PRIOR FILING DATE: 1998-05-07  
120 PRIOR APPLICATION NUMBER: 60/085149  
121 PRIOR FILING DATE: 1998-05-12  
122 PRIOR APPLICATION NUMBER: 60/085323  
123 PRIOR FILING DATE: 1998-05-13  
124 PRIOR APPLICATION NUMBER: 60/085338  
125 PRIOR FILING DATE: 1998-05-13  
126 PRIOR APPLICATION NUMBER: 60/085339  
127 PRIOR FILING DATE: 1998-05-13  
128 PRIOR APPLICATION NUMBER: 60/085579  
129 PRIOR FILING DATE: 1998-05-15  
130 PRIOR APPLICATION NUMBER: 60/085697  
131 PRIOR FILING DATE: 1998-05-15  
132 PRIOR APPLICATION NUMBER: 60/085704  
133 PRIOR FILING DATE: 1998-05-15  
134 PRIOR APPLICATION NUMBER: 60/086414  
135 PRIOR FILING DATE: 1998-05-22  
136 PRIOR APPLICATION NUMBER: 60/086430  
137 PRIOR FILING DATE: 1998-05-22  
138 PRIOR APPLICATION NUMBER: 60/087106  
139 PRIOR FILING DATE: 1998-05-28  
140 PRIOR APPLICATION NUMBER: 60/088026  
141 PRIOR FILING DATE: 1998-06-04  
142 PRIOR APPLICATION NUMBER: 60/088730  
143 PRIOR FILING DATE: 1998-06-10  
144 PRIOR APPLICATION NUMBER: 60/088741  
145 PRIOR FILING DATE: 1998-06-10  
146 PRIOR APPLICATION NUMBER: 60/088810

; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 19/98-06-11
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090538
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091982

Query Match 32.98; Score 561.4; DB 14; Length 1266;
Best Local Similarity 78.08; Pred. No. 4e-165;
Matches 701; Conservative 0; Mismatches 196; Indels 2; Gaps 2;

Qy 243 GGTGACATGAGGGGACCCCACTGATCCATCTTTGGCCACTTCTTCCCTGCTGCTCTC 302
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 303 TCAATGTTGTGCGGAGCTGTGCGGACACCCCTGTACCTGTCTTGGCCACTTCTTCCCTGCTCTC 362
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 64 TCAAGTGTGCTGCGGAGCTGTGCGGACACCCCTGTACCTGTCTTGGCCACTTCTTCCCTGCTCTC 123
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 423 AGGCTGGGGAGTCTGCGGACACCCCTGTACCTGTCTTGGCCACTTCTTCCCTGCTCTC 482
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 184 CGGCTGGGGAGCTGCGGAGCTGTGCGGACACCCCTGTACCTGTCTTGGCCACTTCTTCCCTGCTCTC 243
Qy 483 CAGCCTGGGAGGCTGCGGAGCTGTGCGGACACCCCTGTACCTGTCTTGGCCACTTCTTCCCTGCTCTC 542
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 244 CAGCCTGGGAGGCTGCGGAGCTGTGCGGACACCCCTGTACCTGTCTTGGCCACTTCTTCCCTGCTCTC 303
Qy 543 AGCTGTGAGTGAATGCGGAGCTGTGCGGACACCCCTGTACCTGTCTTGGCCACTTCTTCCCTGCTCTC 602
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 304 AGCTGTGAGTGAATGCGGAGCTGTGCGGACACCCCTGTACCTGTCTTGGCCACTTCTTCCCTGCTCTC 363
Qy 603 GTCTGTGCGGAGCTGCGGAGCTGTGCGGACACCCCTGTACCTGTCTTGGCCACTTCTTCCCTGCTCTC 662
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 364 ATCCGCTGCGGAGCTGCGGAGCTGTGCGGACACCCCTGTACCTGTCTTGGCCACTTCTTCCCTGCTCTC 423
Qy 663 CGGCTGCGGAGCTGCGGAGCTGTGCGGACACCCCTGTACCTGTCTTGGCCACTTCTTCCCTGCTCTC 722
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 424 CGGCTGCGGAGCTGCGGAGCTGTGCGGACACCCCTGTACCTGTCTTGGCCACTTCTTCCCTGCTCTC 483
Qy 723 CCCAGTGGGTATGTGACGAGGAGTGTGCGGACACCCCTGTACCTGTCTTGGCCACTTCTTCCCTGCTCTC 782
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 484 CTTGAGTGGGTATGTGACGAGGAGTGTGCGGACACCCCTGTACCTGTCTTGGCCACTTCTTCCCTGCTCTC 543
Qy 783 CACCAACTTCTGCGGAGCTGTGCGGACACCCCTGTACCTGTCTTGGCCACTTCTTCCCTGCTCTC 842
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 544 CCCCAGTTTCTGCGGAGCTGTGCGGACACCCCTGTACCTGTCTTGGCCACTTCTTCCCTGCTCTC 603
Qy 843 ACAGCCTGGGAGCTGTGCGGAGCTGTGCGGACACCCCTGTACCTGTCTTGGCCACTTCTTCCCTGCTCTC 902
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 604 ACGGCTGGGAGCTGTGCGGACACCCCTGTGCGGATGCGGACACCCCTGCTCTC 663
Qy 903 CAGAACCGATTCTGCCAACTGGAGATCCAAACGCGCTGTGCTGCTGCCAGACCTGCTG 962
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 564 CAGAACCGATTCTGCCAACTGGAGATCCAAACGCGCTGTGCTGCTGCCAGACCTGCTG 723
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 963 GCAGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1021
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 724 CCCTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 783
Qy 1022 ACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1081
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 784 ACGGCTGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 843
Qy 1082 CT-CTTCTCCATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1139
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 844 GTCCGTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 902

RESULT 9
US-10-146-726-319
; Sequence 319, Application US/10146726
; Publication No. US20030129690A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C308
; CURRENT FILING DATE: 2002-05-15
; PRIOR APPLICATION REMOVED - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 319
; LENGTH: 1266
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-146-726-319

Query Match 32.98; Score 561.4; DB 14; Length 1266;
Best Local Similarity 78.08; Pred. No. 4e-165;
Matches 701; Conservative 0; Mismatches 196; Indels 2; Gaps 2;
Qy 243 GGTGACATGAGGGGACCCCACTGATCCATCTTTGGCCACTTCTTCCCTGCTCTC 302
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 4 GGGGACATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 63
Qy 303 TCAATGTTGTGCGGAGCTGTGCGGACACCCCTGTACCTGTCTTGGCCACTTCTTCCCTGCTCTC 362
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 64 TCAAGTGTGCTGCGGAGCTGTGCGGACACCCCTGTACCTGTCTTGGCCACTTCTTCCCTGCTCTC 123
Qy 363 TGCCCAACAGGGGAGTGTGCGGAGCTGTGCGGACACCCCTGTACCTGTCTTGGCCACTTCTTCCCTGCTCTC 422
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 124 TGCCCGCTGCGGAGCTGTGCGGAGCTGTGCGGACACCCCTGTACCTGTCTTGGCCACTTCTTCCCTGCTCTC 183
Qy 423 AGGCTGGGGAGTCTGCGGAGCTGTGCGGACACCCCTGTACCTGTCTTGGCCACTTCTTCCCTGCTCTC 482
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 184 CGGCTGGGGAGGCTGCGGAGCTGTGCGGACACCCCTGTACCTGTCTTGGCCACTTCTTCCCTGCTCTC 243



RESULT 11  
 US-10-146-788-319  
 ; Sequence 319, Application US/10146788  
 ; Publication No. US20030129693A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Baker, Kevin P.  
 ; APPLICANT: Beresini, Maureen  
 ; APPLICANT: DeForge, Laura  
 ; APPLICANT: Desnoyers, Luc  
 ; APPLICANT: Filvaroff, Ellen  
 ; APPLICANT: Gao, Wei-Qiang  
 ; APPLICANT: Gerritsen, Mary E.  
 ; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Godowski, Paul J.  
 ; APPLICANT: Gurney, Austin L.  
 ; APPLICANT: Sherwood, Steven  
 ; APPLICANT: Smith, Victoria  
 ; APPLICANT: Stewart, Timothy A.  
 ; APPLICANT: Tamas, Daniel  
 ; APPLICANT: Watanabe, Colin K  
 ; APPLICANT: Wood, William  
 ; APPLICANT: Zhang, Zemin  
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
 ; FILE REFERENCE: P3330RIC322  
 ; CURRENT APPLICATION NUMBER: US/10/146,788  
 ; CURRENT FILING DATE: 2002-05-15  
 ; Prior Application removed - See File Wrapper or Palm  
 ; NUMBER OF SEQ ID NOS: 550  
 ; SEQ ID NO 319  
 ; LENGTH: 1266  
 ; TYPE: DNA  
 ; ORGANISM: Homo Sapien  
 US-10-146-788-319

Query Match 32.9%; Score 561.4; DB 14; Length 1266;  
 Best Local Similarity 78.0%; Pred. No. 4e-165;  
 Matches 701; Conservative 0; Mismatches 196; Indels 2; Gaps 2;

|    |     |  |     |
|----|-----|--|-----|
| QY | 243 | GGTGACATGAGGGGAGCCACTGCTGGCCACTTCTGGCCACTTCTCTTCTCTGCTGCTCTC | 302 |
| DB | 4   | GGGACATGAGAGGCACACCGACCCCTCTGGCCCTTCTCCCTCTCTGCTGCTCTC       | 63  |
| QY | 303 | TCAATGGTGTGCCAGCTGTGCCGACACCCCTGTACCTGCTCTTGGACACCAACCCAG    | 362 |
| DB | 64  | TCAAAGTGTGCTACCCAGCTGTGCCGACACCACTGTACCTGCTCTGCTGCTGCTGCTG   | 123 |
| QY | 363 | TGCCACAGGGGGTACCCCTGCTGTGATGGCTGTGGCTGCTCTTAAAGTGTGTGACGG    | 422 |
| DB | 124 | TGCCCGCTGGAGTACCCCTGCTGTGATGGCTGTGGCTGCTGCTGCTGCTGCTGCTG     | 183 |
| QY | 423 | AGSGTGGGGAGTCTCGACACCTGATCTGCGACCCCTGCTGCTGCTGCTGCTGCTGCTG   | 482 |
| DB | 184 | CGSGTGGGGAGCCCTGCGACCACTTCCCTGCTGCGACCCCTGCTGCTGCTGCTGCTG    | 243 |
| QY | 483 | CAGCCTGGGGAGCCCTGCGGCGCATGCGGCTGTGTCTCTCTTGGATGAGGATGACGGT   | 542 |
| DB | 244 | CAGCCCGGGGAGACCCGCTGTGCGGGGGCCCTGTGCTCTTGGCAGAGGACGACG       | 303 |
| QY | 543 | AGCTGTAGGTGAATGGCCGAGGTACTCTGATGGAGAGACCTTTAAACCAATTCGAGG    | 602 |
| DB | 304 | AGCTGTAGGTGAACGGCGCTGTATTCGGGAAGGGAGACCTTCCAGCCCACTGACG      | 363 |
| QY | 603 | GTCCCTGCGCGTGTATGACGGTGGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTG     | 662 |
| DB | 364 | ATCCGCTGCGCTGCGAGGAGGGGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTG      | 423 |
| QY | 663 | CGSGTCCCGAGCTGGAGTGCACCGCCCAAGAGAATACAGGTGCCAGGAAGTGTGCTG    | 722 |
| DB | 424 | CGGCTGCCAGCTGGAGTGTCCCCACCCAGGAGGGTCTGAGGTCTTGGCAAGTGTGCTG   | 483 |
| QY | 723 | CCCGATGGGTATGTACCAAGGAGTGTACCCGCGATGCCAGCGCTCCAGCGGCGCAAGGA  | 782 |

US-10-152-380-319  
 ; Sequence 319, Application US/10152380  
 ; Publication No. US20030129694A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Baker, Kevin P.  
 ; APPLICANT: DeForge, Laura  
 ; APPLICANT: Desnoyers, Luc  
 ; APPLICANT: Filvaroff, Ellen  
 ; APPLICANT: Gao, Wei-Qiang  
 ; APPLICANT: Gerritsen, Mary E.  
 ; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Godowski, Paul J.  
 ; APPLICANT: Gurney, Austin L.  
 ; APPLICANT: Sherwood, Steven  
 ; APPLICANT: Smith, Victoria  
 ; APPLICANT: Tamas, Daniel  
 ; APPLICANT: Watanabe, Colin K  
 ; APPLICANT: Wood, William  
 ; APPLICANT: Zhang, Zemin  
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
 ; FILE REFERENCE: P3330RIC397  
 ; CURRENT APPLICATION NUMBER: US/10/152,380  
 ; CURRENT FILING DATE: 2002-05-21  
 ; Prior Application removed - See Palm or File Wrapper  
 ; NUMBER OF SEQ ID NOS: 550  
 ; SEQ ID NO 319  
 ; LENGTH: 1266  
 ; TYPE: DNA  
 ; ORGANISM: Homo Sapien  
 US-10-152-380-319

Query Match 32.9%; Score 561.4; DB 14; Length 1266;  
 Best Local Similarity 78.0%; Pred. No. 4e-165;  
 Matches 701; Conservative 0; Mismatches 196; Indels 2; Gaps 2;

|    |     |  |     |
|----|-----|--|-----|
| QY | 243 | GGTGACATGAGGGGAGCCCACTGCTGCGGACCACTTCTGGCCACTTCTCTTCTCTGCTTCTC | 302 |
| DB | 4   | GGGACATGAGAGGCACACCGACCCCTCTGGCCCTTCTCCCTCTCTGCTGCTCTC         | 63  |
| QY | 303 | TCAATGGTGTGCCAGCTGTGCCGACACCCCTGTACCTGCTCTTGGACACCAACCCAG      | 362 |
| DB | 64  | TCAAAGTGTGCTACCCAGCTGTGCCGACACCACTGTACCTGCTCTGCTGCTGCTGCTG     | 123 |
| QY | 363 | TGCCACAGGGGGTACCCCTGCTGTGATGGCTGTGGCTGCTCTTAAAGTGTGTGACGG      | 422 |
| DB | 124 | TGCCCGCTGGAGTACCCCTGCTGTGATGGCTGTGGCTGCTGCTGCTGCTGCTGCTG       | 183 |
| QY | 423 | AGSGTGGGGAGTCTCGACACCTGATCTGCGACCCCTGCTGCTGCTGCTGCTGCTGCTG     | 482 |
| DB | 184 | CGSGTGGGGAGCCCTGCGACCACTTCCCTGCTGCGACCCCTGCTGCTGCTGCTGCTG      | 243 |
| QY | 483 | CAGCCTGGGGAGCCCTGCGGCGCATGCGGCTGTGTCTCTCTTGGATGAGGATGACGGT     | 542 |
| DB | 244 | CAGCCCGGGGAGACCCGCTGTGCGGGGGCCCTGTGCTCTTGGCAGAGGACGACG         | 303 |
| QY | 543 | AGCTGTAGGTGAATGGCCGAGGTACTCTGATGGAGAGACCTTTAAACCAATTCGAGG      | 602 |
| DB | 304 | AGCTGTAGGTGAACGGCGCTGTATTCGGGAAGGGAGACCTTCCAGCCCACTGACG        | 363 |
| QY | 603 | GTCCCTGCGCGTGTATGACGGTGGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTG       | 662 |
| DB | 364 | ATCCGCTGCGCTGCGAGGAGGGGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTG        | 423 |
| QY | 663 | CGSGTCCCGAGCTGGAGTGCACCGCCCAAGAGAATACAGGTGCCAGGAAGTGTGCTG      | 722 |
| DB | 424 | CGGCTGCCAGCTGGAGTGTCCCCACCCAGGAGGGTCTGAGGTCTTGGCAAGTGTGCTG     | 483 |
| QY | 723 | CCCGATGGGTATGTACCAAGGAGTGTACCCGCGATGCCAGCGCTCCAGCGGCGCAAGGA    | 782 |





Db 784 ACGGTGTCACCATCCAGCTGTGGCCCTGTGCCCTGGGCCCTGGGCTGATGGAAGATG 843  
Qy 1082 CT-CTTCTCCACTCTTGGTGCAGTTAACTGTCTCTGCTGGATTCACTGTGTAGAGC 1139  
Db 844 GTCCGTGCCAGGCCCTGTGCTGCAGGCAACACTTTAGCTTGGTGTCCACCATGCGAAC 902

RESULT 14  
US-10-028-072-319  
; Sequence 319, Application US/10028072  
; Publication No. US20030004311A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Collin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang  
; TITLE OF INVENTION:  
; FILE REFERENCE:  
; CURRENT APPLICATION NUMBER: US/10/028,072  
; CURRENT FILING DATE: 2001-12-19  
; PRIOR APPLICATION NUMBER: 60/049911  
; PRIOR FILING DATE: 1997-06-18  
; PRIOR APPLICATION NUMBER: 60/036974  
; PRIOR FILING DATE: 1997-08-26  
; PRIOR APPLICATION NUMBER: 60/059113  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/059115  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/059117  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/059122  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/059184  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/059263  
; PRIOR FILING DATE: 1997-09-18  
; PRIOR APPLICATION NUMBER: 60/059352  
; PRIOR FILING DATE: 1997-09-19  
; PRIOR APPLICATION NUMBER: 60/059588  
; PRIOR FILING DATE: 1997-09-19  
; PRIOR APPLICATION NUMBER: 60/059836  
; PRIOR FILING DATE: 1997-09-24  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/062285  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/062287  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/062814  
; PRIOR FILING DATE: 1997-10-24  
; PRIOR APPLICATION NUMBER: 60/062816  
; PRIOR FILING DATE: 1997-10-24  
; PRIOR APPLICATION NUMBER: 60/063045  
; PRIOR FILING DATE: 1997-10-24  
; PRIOR APPLICATION NUMBER: 60/063082  
; PRIOR FILING DATE: 1997-10-31  
; PRIOR APPLICATION NUMBER: 60/063127  
; PRIOR FILING DATE: 1997-10-24  
; PRIOR APPLICATION NUMBER: 60/063327  
; PRIOR FILING DATE: 1997-10-27  
; PRIOR APPLICATION NUMBER: 60/063329  
; PRIOR FILING DATE: 1997-10-27  
; PRIOR APPLICATION NUMBER: 60/063550  
; PRIOR FILING DATE: 1997-10-28  
; PRIOR APPLICATION NUMBER: 60/063561  
; PRIOR FILING DATE: 1997-10-28  
; PRIOR APPLICATION NUMBER: 60/063704  
; PRIOR FILING DATE: 1997-10-29  
; PRIOR APPLICATION NUMBER: 60/063733  
; PRIOR FILING DATE: 1997-10-29  
; PRIOR APPLICATION NUMBER: 60/063735  
; PRIOR FILING DATE: 1997-10-29  
; PRIOR APPLICATION NUMBER: 60/063738  
; PRIOR FILING DATE: 1997-10-29  
; PRIOR APPLICATION NUMBER: 60/063755  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/064248  
; PRIOR FILING DATE: 1997-11-03  
; PRIOR APPLICATION NUMBER: 60/064809  
; PRIOR FILING DATE: 1997-11-07  
; PRIOR APPLICATION NUMBER: 60/065186  
; PRIOR FILING DATE: 1997-11-12  
; PRIOR APPLICATION NUMBER: 60/065846  
; PRIOR FILING DATE: 1997-11-17  
; PRIOR APPLICATION NUMBER: 60/066364  
; PRIOR FILING DATE: 1997-11-21  
; PRIOR APPLICATION NUMBER: 60/066453  
; PRIOR FILING DATE: 1997-11-24  
; PRIOR APPLICATION NUMBER: 60/066511  
; PRIOR FILING DATE: 1997-11-24  
; PRIOR APPLICATION NUMBER: 60/066770  
; PRIOR FILING DATE: 1997-11-24  
; PRIOR APPLICATION NUMBER: 60/069212  
; PRIOR FILING DATE: 1997-12-11  
; PRIOR APPLICATION NUMBER: 60/069278  
; PRIOR FILING DATE: 1997-12-11  
; PRIOR APPLICATION NUMBER: 60/069334  
; PRIOR FILING DATE: 1997-12-11  
; PRIOR APPLICATION NUMBER: 60/069694  
; PRIOR FILING DATE: 1997-12-16  
; PRIOR APPLICATION NUMBER: 60/072320  
; PRIOR FILING DATE: 1998-01-23  
; PRIOR APPLICATION NUMBER: 60/073612  
; PRIOR FILING DATE: 1998-02-04  
; PRIOR APPLICATION NUMBER: 60/074086  
; PRIOR FILING DATE: 1998-02-09  
; PRIOR APPLICATION NUMBER: 60/074092  
; PRIOR FILING DATE: 1998-02-09  
; PRIOR APPLICATION NUMBER: 60/077791  
; PRIOR FILING DATE: 1998-03-12  
; PRIOR APPLICATION NUMBER: 60/078910  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/079294  
; PRIOR FILING DATE: 1998-03-25  
; PRIOR APPLICATION NUMBER: 60/079663  
; PRIOR FILING DATE: 1998-02-27  
; PRIOR APPLICATION NUMBER: 60/079728  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/080165  
; PRIOR FILING DATE: 1998-03-31  
; PRIOR APPLICATION NUMBER: 60/081203  
; PRIOR FILING DATE: 1998-04-09  
; PRIOR APPLICATION NUMBER: 60/081229  
; PRIOR FILING DATE: 1998-04-09  
; PRIOR APPLICATION NUMBER: 60/081695  
; PRIOR FILING DATE: 1998-04-14  
; PRIOR APPLICATION NUMBER: 60/081817  
; PRIOR FILING DATE: 1998-04-15  
; PRIOR APPLICATION NUMBER: 60/081818  
; PRIOR FILING DATE: 1998-04-15  
; PRIOR APPLICATION NUMBER: 60/082999  
; PRIOR FILING DATE: 1998-04-24

; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085149
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/086414
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/086430
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088730
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088741
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 19/98-06-11
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090538
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07

Query Match 32.9%; Score 561.4; DB 15; Length 1266;
Best Local Similarity 78.0%; Pred. No. 4e-165;
Matches 701; Conservative 0; Mismatches 196; Indels 2; Gaps 2;
QY 243 GGTGACATGAGGGGCGCCACTGATCCATCTTCTGGCCACTTCTCTCTGCTCTCTC 302
DB 4 GGGGACATGAGAGGACACCGGAGACCCACCTCTCTGGGCTCTCTCTCTCTCTCTC 53

QY 303 TCAATGGTGTGTGCCCCAGCTGTGCCGACACCCCTGTACCTGTCTTGGACACCCACCCAG 362
DB 64 TCAAGGTGCGTACCACCTGTGCCGACACCATGTACCTGTCCCTGCCACCTCCCGCA 123
QY 363 TCCCCACAGGGGTACCCTGTGTGGTGTGGTGTGGTGTAAAGTGTGTGTGTCAGGG 422
DB 124 TGCCCGCTGGGAGTACCCTGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 183
QY 423 AGGCTGGGGAGTCCCTGTGGACACCTGTCTCGACCCACCCAGCCAGCGCTGTGTGT 482
DB 184 CGCTGGGGAGCCCTGTGGACCACTCCACGTCTCGAGCCAGCCAGCGCTGTGTGTG 243
QY 483 CAGCCTGGGGAGCCCTGTGGCCCATGGGCTGTGTCTCTTGGATGAGATGACGGT 542
DB 244 CAGCCCGGGAGGACCCGTGGCCGGGGCCCTGTGCTCTTGGCAGAGGACGACAGC 303
QY 543 AGCTGTGAGTGAATGGCCGACGATACCTGTGATGAGAGACCTTTAAACCCANTTGCAGG 602
DB 304 AGCTGTGAGTGAACGGCCGCTGTATCGGGAAGGGAGACCTTCCAGCCCACTGCGAGC 363
QY 603 GTCCCTGTGCGCTGTGATGAGGTGGCTTACCTGCTCCCTGTGCTGCTGAGTGTG 662
DB 364 ATCCGCTGCGCTGTGAGGACGGCGCTTACCTGCTGCGCTGTGCGAGGAGTGTG 423
QY 663 CGGCTGCCAGCTGGGACTGCCACGCCCAAGAGATACAGGTGCCAGGAAAGTGTCTGC 722
DB 424 CGGCTGCCAGCTGGGACTGCCCCACCCAGAGGGTCTGAGGTCTGCGGCAAGTGTCTGC 483
QY 723 CCCGAGTGGTATGTACACGAGGAGTGACACGGGGATCCAGCGCTCCACGCGGCAAGGA 782
DB 484 CTGAGTGGGTGTGCGCCAAAGGAGGGGACTTGGGGACCCAGCCCTTCCAGCCCAAGGA 543
QY 783 CACCAACTTCTGCGCTTGTCTACTCTGCTGTGATGCTCTTGTCCAAATTTGGAGC 842
DB 544 CCCAGATTTCTGGCTTGTCTCTTCCCTGCCCCCTGTGCTCTGCCCCAGATGGAGC 603
QY 843 ACAGCCTGGGCCCCCTGCTCAACACCTGTGGGCTGGGCTAGGACACCCAGTGTCCAAC 902
DB 604 ACGGCTGGGACCTGCTCGACACCTGTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 663
QY 903 CAGAACCGATCTGCCAACTGGAGATCCAAACCGGCTGTGTCTGCCAGACCCCTGCCTG 962
DB 664 CAGAACCGCTTCTGCGGACTGGAGACCCAGGCGCCCTGTGCTGTCCAGGCCCCGCGCA 723
QY 963 GCAGCCAGGAGCCAGCTCATGGAACAGTGTCTTA-AGSCCAACTGGGATCGGAT 1021
DB 724 CCCTCCAGGGTGGAGTCCACAAACAGTGCCTTCTAGAGCCGGCTGGGAATGGGAC 783
QY 1022 ACAGGCTGCGCATCTCAGCAATGACCCCTAGGACACGAGGCGCTGAGTGTGTAGTG 1081
DB 784 ACGGTGTCCACCATGCCAGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGATG 843
QY 1082 CT-CTTCTCCATGCTTGGCTGCACTTAACCTGCTCTCTGATTCACCTGTGTAGAGC 1139
DB 844 GTCCGTGCCAGGCGCTTGGCTGCAAGCAACACTTTAGCTTGGGTCCACCATGCGAGAAC 902

RESULT 15
US-10-121-049-319
; Sequence 319, Application US/10121049
; Publication No. US2003002239A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 18:02:17 ; Search time 3028.68 Seconds

(without alignments)  
9133.322 Million cell updates/sec

Title: US-10-010-408-1

Perfect score: 1708

Sequence: 1 GACGCTTCTGATCCAGAG.....GCCTAGATAACACCCAAA 1708

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*\*

2: em\_esthum:\*\*

3: em\_estin:\*\*

4: em\_estmu:\*\*

5: em\_estov:\*\*

6: em\_estopl:\*\*

7: em\_estro:\*\*

8: em\_hic:\*\*

9: gb\_est1:\*\*

10: gb\_est2:\*\*

11: gb\_hic:\*\*

12: gb\_est3:\*\*

13: gb\_est4:\*\*

14: gb\_est5:\*\*

15: em\_estfun:\*\*

16: em\_estom:\*\*

17: gb\_gss:\*\*

18: em\_gss\_hum:\*\*

19: em\_gss\_inv:\*\*

20: em\_gss\_pln:\*\*

21: em\_gss\_vrt:\*\*

22: em\_gss\_fun:\*\*

23: em\_gss\_mam:\*\*

24: em\_gss\_mus:\*\*

25: em\_gss\_Other:\*\*

26: em\_gss\_pro:\*\*

27: em\_gss\_rod:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query Match | Length | ID   | Description |
|------------|-------|-------------|--------|------|-------------|
| c          | 1     | 725.4       | 42.5   | 762  | 14 BQ195526 |
|            | 2     | 511.2       | 29.9   | 1058 | 14 BM805088 |
|            | 3     | 501.8       | 29.4   | 979  | 14 BQ279131 |
|            | 4     | 494.6       | 29.0   | 886  | 13 BI822142 |
|            | 5     | 440.6       | 25.8   | 916  | 13 BI457367 |
|            | 6     | 436.4       | 25.6   | 750  | 13 BM043988 |

|    |       |      |      |    |          |          |             |
|----|-------|------|------|----|----------|----------|-------------|
| 7  | 432.4 | 25.3 | 792  | 12 | BF138093 | BF138093 | 6017850003  |
| 8  | 430.2 | 25.2 | 651  | 13 | BI457141 | BI457141 | 6031853392  |
| 9  | 429.4 | 25.1 | 966  | 12 | BF141695 | BF141695 | 6017907552  |
| 10 | 425.8 | 24.9 | 940  | 14 | BQ937887 | BQ937887 | AGENCYCOURT |
| 11 | 421.2 | 24.7 | 1006 | 14 | BM921531 | BM921531 | AGENCYCOURT |
| 12 | 417.8 | 24.5 | 1073 | 14 | BQ073722 | BQ073722 | AGENCYCOURT |
| 13 | 411.6 | 24.1 | 790  | 13 | BM046275 | BM046275 | 603626068   |
| 14 | 410.2 | 24.0 | 800  | 13 | BI826781 | BI826781 | 603077268   |
| 15 | 408.2 | 23.9 | 928  | 13 | BI161474 | BI161474 | 602864871   |
| 16 | 399.2 | 23.4 | 888  | 13 | BI825652 | BI825652 | 603072631   |
| 17 | 385.2 | 22.6 | 537  | 14 | BQ560868 | BQ560868 | H4067A01-   |
| 18 | 375.2 | 22.0 | 1022 | 14 | BQ952960 | BQ952960 | AGENCYCOURT |
| 19 | 366   | 21.4 | 620  | 13 | BQ928868 | BQ928868 | HNC57-1-D   |
| 20 | 361.2 | 21.1 | 1166 | 13 | BM543799 | BM543799 | AGENCYCOURT |
| 21 | 349.6 | 20.5 | 749  | 9  | AL555144 | AL555144 | AL555144    |
| 22 | 346.8 | 20.3 | 488  | 9  | AA754979 | AA754979 | vu55908.r   |
| 23 | 336.4 | 19.7 | 1251 | 14 | BQ961357 | BQ961357 | AGENCYCOURT |
| 24 | 335.8 | 19.7 | 424  | 10 | BB849097 | BB849097 | BB849097    |
| 25 | 321.6 | 18.8 | 489  | 14 | BM751866 | BM751866 | K-EST0028   |
| 26 | 318.4 | 18.6 | 426  | 9  | AA717584 | AA717584 | vt997c08.r  |
| 27 | 304.6 | 17.8 | 452  | 10 | BE481184 | BE481184 | 166529 BA   |
| 28 | 297.8 | 17.4 | 933  | 14 | BQ278961 | BQ278961 | AGENCYCOURT |
| 29 | 234.4 | 13.7 | 657  | 13 | BM488499 | BM488499 | pgm2p.pk0   |
| 30 | 231.4 | 13.5 | 418  | 9  | AI225477 | AI225477 | ue88D01.y   |
| 31 | 226.4 | 13.3 | 618  | 12 | BG538695 | BG538695 | 602566932   |
| 32 | 223.6 | 13.1 | 792  | 13 | BI823598 | BI823598 | 603040962   |
| 33 | 223.2 | 13.1 | 405  | 12 | BG900069 | BG900069 | HOA51-1-A   |
| 34 | 209.6 | 12.3 | 380  | 12 | BG900020 | BG900020 | HOA48-1-G   |
| 35 | 209.4 | 12.3 | 823  | 14 | BQ067127 | BQ067127 | AGENCYCOURT |
| 36 | 209.2 | 12.2 | 668  | 9  | AL555143 | AL555143 | AL555143    |
| 37 | 199.4 | 11.7 | 401  | 12 | BF849398 | BF849398 | CM3-EN007   |
| 38 | 180   | 10.5 | 369  | 10 | BB792544 | BB792544 | BB792544    |
| 39 | 179.2 | 10.5 | 345  | 10 | BB222795 | BB222795 | BB222795    |
| 40 | 178.4 | 10.4 | 337  | 10 | BB319151 | BB319151 | BB319151    |
| 41 | 169.4 | 9.9  | 307  | 10 | BB220676 | BB220676 | BB220676    |
| 42 | 167.2 | 9.8  | 313  | 10 | BB367824 | BB367824 | BB367824    |
| 43 | 164   | 9.6  | 315  | 10 | BB374499 | BB374499 | BB374499    |
| 44 | 161.8 | 9.5  | 697  | 12 | BF056642 | BF056642 | 7k20f09.x   |
| 45 | 154.4 | 9.0  | 948  | 9  | AL542806 | AL542806 | AL542806    |

## ALIGNMENTS

|            |   |             |                   |            |        |                 |
|------------|---|-------------|-------------------|------------|--------|-----------------|
| RESULT 1   | BQ195526/c  | BQ195526    | 762 bp            | mRNA       | linear | EST 30-APR-2002 |
| LOCUS      | UI-R-CN1-CMq-k-07-0-UI.s1   | UI-R-CN1    | Rattus norvegicus | CDNA clone |        |                 |
| DEFINITION | UI-R-CN1-CMq-k-07-0-UI 3', mRNA sequence.   |             |                   |            |        |                 |
| ACCESSION  | BQ195526  |             |                   |            |        |                 |
| VERSION    | BQ195526.1  | GI:20371077 |                   |            |        |                 |
| KEYWORDS   | EST   |             |                   |            |        |                 |
| SOURCE     | Norway rat.   |             |                   |            |        |                 |
| ORGANISM   | Rattus norvegicus   |             |                   |            |        |                 |
| REFERENCE  | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.   |             |                   |            |        |                 |
| AUTHORS    | 1 (bases 1 to 762)  |             |                   |            |        |                 |
| TITLE      | Bonaldo,M.F., Lennon,G. and Soares,M.B.   |             |                   |            |        |                 |
| JOURNAL    | Normalization and subtraction: two approaches to facilitate gene discovery  |             |                   |            |        |                 |
| MEDLINE    | Genome Res. 6 (9), 791-806 (1996)   |             |                   |            |        |                 |
| COMMENT    | 97044477  |             |                   |            |        |                 |
|            | Contact: Soares, MB   |             |                   |            |        |                 |
|            | Program for Rat Gene Discovery and Mapping  |             |                   |            |        |                 |
|            | University of Iowa  |             |                   |            |        |                 |
|            | 451 Eckstein Medical  |             |                   |            |        |                 |
|            | Tel: 319 335 8250   |             |                   |            |        |                 |
|            | Fax: 319 335 9565   |             |                   |            |        |                 |
|            | Email: msoares@blue.weeg.uiowa.edu  |             |                   |            |        |                 |
|            | The sequence contained an oligo-dr track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A |             |                   |            |        |                 |

tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to identify it as a clone from the normalized cervix library cDNA Library Preparation: M.B. Soares Lab Clone distribution: clones will be available through Research Genetics (www.resgen.com)  
Seq primer: M13 Forward  
POLYA=yes.

# FEATURES

## Location/Qualifiers

1. .762  
/organism="Rattus norvegicus"  
/strain="Sprague-Dawley"  
/db\_xref="taxon:10116"  
/clone="UI-R-CNI-cmq-k-07-0-UI"  
/clone\_lib="UI-R-CNI"  
/dev\_stage="adult"  
/lab\_host="PH10B (Life Technologies)"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site1: Not I; Site2: Eco RI; The UI-R-CNI library is a subtracted library derived from the following pool of seven normalized rat libraries: normalized rat seminal vesicles, normalized rat penis, normalized rat bladder, normalized rat cervix, normalized rat brown adipose, normalized rat fundus, and normalized rat salivary gland. It was constructed according to the procedure described by Bonaldo, Lennon & Soares (Genome Research Genome 6: 791-806, 1996). For construction of the CNI library, plasmid DNA from the pool of normalized libraries was electroporated into competent bacteria for the production of single-stranded circular DNA. This was then used as a tracer in a subtractive hybridization with a driver (PCR amplified inserts from a plasmid DNA template preparation) comprising: a) a pool of about 34,000 clones from the Rat Unigene Set corresponding to plates R-5-AA-NN excluding plates R-5-MM and MN. This pool represented 40% of the final driver population. b) a pool of about 29,000 clones from subtracted libraries CA0 and CAL corresponding to plates R-CA0-AW through R-CA0-AXS, R-CA0-AZX through R-CA0-BAZ, R-CA0-BFE through R-CA0-BHY, R-CA0-BJS, R-CA0-BKE, R-CA0-BKG-H, R-CA0-BKJ-K, R-CA0-BKP through R-CA0-BKS, R-CA0-BKU-V, R-CA0-BLY through R-CA0-BMA, R-CA0-BMC through R-CA0-BME, R-CA0-BNS, R-CA0-BOB through R-CA0-BOJ, R-CA0-BPA through R-CA0-BPG, R-CAL-BBA through R-CAL-BDA, R-CAL-BHZ through R-CAL-BJF, R-CAL-BJK, R-CAL-BJT through R-CAL-BKB, R-CAL-BKD, R-CAL-BKF, R-CAL-BKI, R-CAL-BKT, R-CAL-BLF, R-CAL-BLH through R-CAL-BLN, R-CAL-BLS, R-CAL-BLU-V, R-CAL-BNR, and R-CAL-BLE. The resulting pool represented 20% of the final driver population. c) a pool of about 15,000 clones from non-normalized libraries CS0s, CT0s, CU0s, CW0s, CX0s and normalized libraries CS0, CT0, CU0, CW0, and CX0 corresponding to plates R-CS0s-CBD through R-CS0s-CBO, R-CT0s-CAM through R-CT0s-CAX, R-CU0s-CBP through R-CU0s-CCA, R-CW0s-CCB through R-CW0s-CCM, R-CX0s-CCN through R-CX0s-CCX, R-CS0-BSD, R-CS0-BTD through R-CS0-BTV, R-CS0-BVM, R-CT0-BTW through R-CT0-BUP, R-CT0-BVN, R-CU0-BUQ through R-CU0-BVL, R-CW0-BVU through R-CW0-BWP, R-CW0-BXN through R-CW0-BXO, R-CX0-BWQ through R-CX0-BXM. The resulting pool represented 5% of the final driver population. d) a pool of about 5,000 clones (1,000 from non-normalized eye library CV0 and 4,000 from normalized eye library CV1) corresponding to plates R-CV0-BRH through R-CV0-BRR, R-CV1-BRS through R-CV1-BSC, R-CV1-BSE through R-CV1-BTC, and R-CV1-BVO through R-CV1-BVD. This pool represented about 5% of the final driver population. e) A pool of about 10,000 clones from subtracted library BS2, BV0 and BV0p (7-9.5 kb cDNA library fraction from rat whole embryo), and BX0 (0.5-7kb cDNA library fraction from rat whole embryo) corresponding to plates R-BS2-BDB through R-BS2-BOF, R-BV0-ANK through R-BV0-ANR, R-BV0p-AOI through R-BV0p-AFB, and R-BX0-AQY through R-BX0-ASH. The resulting pool represented 5% of the final driver population. f) a pool of about 7,000 clones from the seven non-normalized libraries that make up the tracer

including CY0, CZ0, DA0, DB0, DC0, DD0, and DE0 corresponding to plates R-CY0-BXP through R-CY0-BXZ, R-CZ0-BYA through R-CZ0-BYI, R-CZ0-BZB-C, R-DA0-BYJ through R-DA0-BYP, R-DA0-BZD through R-DA0-BZH, R-DB0-BYQ through R-DB0-BZA, R-DC0-BZI through R-DC0-BZQ, R-DD0-CAY through R-DD0-CBA, R-DD0-BZR through R-DD0-CAA, R-DD0-CBB-C, and R-DE0-CAB through R-DE0-CAL. The resulting pool represented about 10% of the final driver population. g) a pool of about 2,000 clones from the pool of normalized libraries, CNO, that makes up the tracer. The corresponding plates are R-CNO-BKW through R-CNO-BLD, R-CNO-BLG, R-CNO-BLP through R-CNO-BLR, R-CNO-BLT, R-CNO-BLW-X, R-CNO-BMB, and R-CNO-BMF through R-CNO-BML. This pool represented 5% of the final driver population. h) a pool of the 28 most abundant clones in the CNO pool corresponding to the following addresses: bkw-a-09-0-UI, bkw-b-09-0-UI, bkw-b-11-0-UI, bkw-b-10-0-UI, bkw-d-01-0-UI, bkw-d-06-0-UI, bkw-g-08-0-UI, bkw-h-12-0-UI, bky-a-05-0-UI, bkz-a-06-0-UI, bkz-a-11-0-UI, bkz-c-06-0-UI, bkz-c-09-0-UI, bkz-d-10-0-UI, bla-a-01-0-UI, bla-a-02-0-UI, bla-f-04-0-UI, bla-g-07-0-UI, bla-g-12-0-UI, blb-a-12-0-UI, blb-f-02-0-UI, blc-a-11-0-UI, blc-a-95-0-UI, blb-1-08-0-UI, blb-f-02-0-UI, blq-h-04-0-UI, blr-a-05-0-UI, blt-f-08-0-UI. This pool represented 5% of the final driver population. i) One abundant CNO clone (corresponding to the address bkz-a-11-0-UI) was digested with Not I and Eco RI and the resulting insert was gel purified. This purified insert was added directly to the driver so that it represented 5% of the final driver population.

TAG\_LTB-UI-R-CNI  
TAG\_TISSUE-cervix  
TAG\_SEQ-GACCA"

BASE COUNT 177 a 176 c 198 g 211 t  
ORIGIN

|                       |                 |                     |           |             |
|-----------------------|-----------------|---------------------|-----------|-------------|
| Query Match           | 42.5%;          | Score 725.4;        | DB 14;    | Length 762; |
| Best Local Similarity | 99.2%;          | Pred. No. 3.9e-186; |           |             |
| Matches 729;          | Conservative 0; | Mismatches 6;       | Indels 0; | Gaps 0;     |

  

|    |      |   |      |
|----|------|---|------|
| QY | 974  | CCACAGCTCATGACAGTGCTTTCTAAGGCCAACTGGGGATCGCGATACAGGCTGCC      | 1033 |
| DB | 762  | CCACAGCTCATGACAGTGCTTTCTAAGGCCAACTGGGGATCGCGATACAGGCTGCC      | 703  |
| QY | 1034 | ATCCTCAGCAATGACCCCTAGGACAGGCCCTGGAGCTGCTGTAGATGCTCTTCTCCATG   | 1093 |
| DB | 702  | ATCCTCAGCAATGACCCCTAGGACAGGCCCTGGAGCTGCTGTAGATGCTCTTCTCCATG   | 643  |
| QY | 1094 | CTCTTGGCTGAGTAACTGTCCTGTTGGATTCACTGTGTAGAGCACTGAGCGATCCC      | 1153 |
| DB | 642  | CTCTTGGCTGAGTAACTGTCCTGTTGGATTCACTGTGTAGAGCACTGAGCGATCCC      | 583  |
| QY | 1154 | TCCTCTGCTGAGTAACTGTCCTGTTGGATTCACTGTGTAGAGCACTGAGCGATCCC      | 1213 |
| DB | 582  | TCCTCTGCTGAGTAACTGTCCTGTTGGATTCACTGTGTAGAGCACTGAGCGATCCC      | 523  |
| QY | 1214 | TTCTGGTGTCTCTGCTCATTCTCAAAACATCCCTGTACAAAAGGACACCAAAAG        | 1273 |
| DB | 522  | TTCTGGTGTCTCTGCTCATTCTCAAAACATCCCTGTACAAAAGGACACCAAAAG        | 463  |
| QY | 1274 | ACCTTTAACTTAGGTATATCTAGGCAACCTGGCCACCGTCTGGGGATAGGTCATG       | 1333 |
| DB | 462  | ACCTTTAACTTAGGTATATCTAGGCAACCTGGCCACCGTCTGGGGATAGGTCATG       | 403  |
| QY | 1334 | TTAGGACACAGACAGATGCTGCTGAACTTCCCAATTCCTTCTGGACTTCTGTATGCT     | 1393 |
| DB | 402  | TTAGGACACAGACAGATGCTGCTGAACTTCCCAATTCCTTCTGGACTTCTGTATGCT     | 343  |
| QY | 1394 | TGTCCTCCCAAGATGATGAATGAACCTCGTAAGTGTACCTTCCCTGACCTGAGAACCCCTG | 1453 |
| DB | 342  | TGTCCTCCCAAGATGATGAATGAACCTCGTAAGTGTACCTTCCCTGACCTGAGAACCCCTG | 283  |
| QY | 1454 | CCTGCTCGGAGATATTACAGGGCAGAGATTCCTCTGTGAACATGAGAGATGAATCACT    | 1513 |

```

Db      282 CCTGCTCGGAAGTATTTCAGGGCAGAAATCTCTGTGAACATGAAGAGATGAATCACACT 223
Qy      1514 GTCTTAAGAAATTCCTGAAGTCCAGGAACCTTGAGCTTTGTATTTTCAGGAATGCACAT 1573
Db      222 GTCTTAAGAAATTCCTGAAGTCCAGGAACCTTGAGCTTTGTATTTTCAGGAATGCACAT 163
Qy      1574 CTCTTAAGCACTCGCAAAACAGGAAGCTCCACACCTCTTGGCAGGCCAGGCCCTTCTCT 1633
Db      162 CTCTTAAGCACTCGCAAAACAGGAAGCTCCACACCTCTTGGCAGGCCAGGCCCTTCTCT 103
Qy      1634 TCAGCATGAGAAACAGGAGGACAGCAGAGTACTCTCTCTGAGGACTAGTCTAGCCTA 1693
Db      102 TCAGCATGAGAAACAGGAGGACAGCAGAGTACTCTCTCTGAGGACTAGTCTAGCCTA 43
Qy      1694 GAATAAACACCCCAA 1708
Db      42 GAATAAACACCCCAA 28

RESULT 2
BM805088
LOCUS      BM805088
DEFINITION AGENCOURT_6490429 NIH_MGC_125 Homo sapiens cDNA clone IMAGE:5725937
            5', mRNA sequence.
ACCESSION  BM805088
VERSION     BM805088.1 GI:19121911
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1058)
            NIH-MGC http://mgi.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgabbs-r@mail.nih.gov
            Tissue Procurement: Invitrogen
            cDNA Library Preparation: Life Technologies, Inc.
            CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12717 row: j column: 18
            High quality sequence stop: 697.
FEATURES
    source
        1..1058
            Location/Qualifiers
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:5725937"
                /clone_lib="NIH_MGC_125"
                /lab_host="DH10B"
                /note="Organ: ovary (pool of 3); Vector: pCMV-SPORT6;
                Site_1: EcoRV (destroyed); Site_2: NotI; RNA source pool
                of three ovaries, from females ranging in age from 38 to
                49 yo. Library is oligo-dT primed and directionally cloned
                (EcoRV site is destroyed upon cloning). Average insert
                size 2.1 kb, insert size range 1-3.5 kb. Library is
                normalized and enriched for full-length clones and was
                constructed by C. Gruber (Invitrogen). Research Genetics
                tracking code 036."
                BASE COUNT  139 a 423 c 294 g 200 t
                ORIGIN
                    Query Match      29.9%; Score 511.2; DB 14; Length 1058;
                    Best local Similarity 79.3%; Pred. No. 8.1e-128;
                    Matches 606; Conservative 0; Mismatches 158; Indels 0; Gaps 0;

Qy      242 CGGTGATCAGGGGAGCCCACTGATCCATCTTCGGCCACTTCCTCTCTGCTCTCT 301
Db      2 CAGGGGAATGAGAGGACACCCGAAGACCCACTCTCGGCTTCTCCCTCTCTCTCTCT 61

```

```

Qy      302 CTCAATGGTGTGTGCCAGCTGTGCCGAGACACACCTGTACCTGTCTTGGACACACCCCA 361
Db      62 CTCAAGGTGTCGTACCAAGCTGTGCCGAGACACCATGTACCTGCCCTGGCCACTCCCG 121
Qy      362 GTGCCACACAGGGGTACCCCTGTGTGTGATGGCTGTGGCTGTGTAAATGTGTGACG 421
Db      122 ATGCCCGCTGGAGTACCCCTGTGTGTGATGGCTGTGGCTGTGTGCGGGTATGTGACG 181
Qy      422 GAGCTGGGGAGTCTCTGCCACACCTGTGATGTCTGCGACCCCGAGCCAGGCGCTGGTTG 481
Db      182 GCGCTGGGGAGCCCTGTGCCAACACTCCACGTCTGCGACCGCCAGCCAGGCGCTGGTCTG 241
Qy      482 TCAGCTGTGGGAGGCGCTGTGCCCATGGCGGTGTGTCTCTTGGATGAGGATGACGG 541
Db      242 CCAGCCCGGGGAGGACCCGCTGGACGGGGGCCCTGTGCTCTTGGCAGAGGACGACAG 301
Qy      542 TAGCTGTGAGGTGAATGGCGCAGGTACCTGTGGATGGAGAGACCTTTAAACCAATTCGAG 601
Db      302 CAGCTGTGAGGTGAACGGCGCCTGTATCGGGAAGGGGAGGACCTTCCAGCCCCACTGCAG 361
Qy      602 GGTCTGTGCGCTGTGATGACGCTGTGCTTCACTGCTGCTGCGCTGTGACGAGGATGT 661
Db      362 CATCCGCTGCGCTGTGAGGAGCGCGCTTCACTTGCCTGCGCTGTGACGAGGATGT 421
Qy      662 GCGGCTGCCAGCTGGGACTGCCACCGCCCAAGAGAATACAGGTGCCAGGAAGTGTCTG 721
Db      422 GCGGCTGCCAGCTGGGACTGCCACCGCCCAAGAGGAGGTGAGGTCTTGGCAAGTGTCTG 481
Qy      722 CCCGAGTGGGTATGTACAGGAGGTGACACCGCGCATCCAGGCTTCCAGGCGCAAGG 781
Db      482 CCCTGATGGGTGTGCGGCAAGGAGGGGACTGGGACCCAGCCCTTCCAGCCCCAAGG 541
Qy      782 ACACCAACTTTCTGCCCTTCTCACTCTCTGCTGTGATGCTCTTGTCCAAATTCGAG 841
Db      542 ACCCAGTTTCTGGCCCTTGTCTTCCCTGCCCTTGTGCTGCTGCCAGATGGAG 601
Qy      842 CACAGCTGGGGCCCTTGTCTCAACCACTGTGGGCTGGGCATAGCCAGCCAGTGTCCAA 901
Db      602 CAGCGCTGGGGACCCCTGTCTCGACCACTGTGGGCTGGGCATAGCCAGCCAGTGTCCAA 661
Qy      902 CGAGACCGATTTCTGCCAATCGAGATCCAAACCGCGCTGTGTCTGCCAGACAGCCCTGCT 961
Db      662 CCAGAACCGCTTCTGCCGACTGGAGACCGCGCTGTGTCTGCCAGCCCTGCTGCC 721
Qy      962 GGCAGCCAGGAGCCACAGCTCATGGAACAGTGTCTTCTAAGGCC 1005
Db      722 ACCCTCCAGGGGTGCGAGTCCACAAACAGTGTCTTCTAGAGCC 765

```

```

RESULT 3
BM805088
LOCUS      BM805088
DEFINITION AGENCOURT_7046721 NIH_MGC_107 Homo sapiens cDNA clone IMAGE:5805819
            5', mRNA sequence.
ACCESSION  BM805088
VERSION     BM805088.1 GI:20489339
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            NIH-MGC http://mgi.nci.nih.gov/.
            1 (bases 1 to 979)
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgabbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            CDNA Library Preparation: Rubin Laboratory
            CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12717 row: j column: 18
            High quality sequence stop: 697.
FEATURES
    source
        1..1058
            Location/Qualifiers
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:5725937"
                /clone_lib="NIH_MGC_125"
                /lab_host="DH10B"
                /note="Organ: ovary (pool of 3); Vector: pCMV-SPORT6;
                Site_1: EcoRV (destroyed); Site_2: NotI; RNA source pool
                of three ovaries, from females ranging in age from 38 to
                49 yo. Library is oligo-dT primed and directionally cloned
                (EcoRV site is destroyed upon cloning). Average insert
                size 2.1 kb, insert size range 1-3.5 kb. Library is
                normalized and enriched for full-length clones and was
                constructed by C. Gruber (Invitrogen). Research Genetics
                tracking code 036."
                BASE COUNT  139 a 423 c 294 g 200 t
                ORIGIN
                    Query Match      29.9%; Score 511.2; DB 14; Length 1058;
                    Best local Similarity 79.3%; Pred. No. 8.1e-128;
                    Matches 606; Conservative 0; Mismatches 158; Indels 0; Gaps 0;

Qy      242 CGGTGATCAGGGGAGCCCACTGATCCATCTTCGGCCACTTCCTCTCTGCTCTCT 301
Db      2 CAGGGGAATGAGAGGACACCCGAAGACCCACTCTCGGCTTCTCCCTCTCTCTCTCT 61

```







|         | BASE COUNT | ORIGIN |
|---------|------------|--------|
| 193 a   | 204 g      | 190 e  |
| 193 b   | 205 g      | 190 e  |
| 193 c   | 206 g      | 190 e  |
| 193 d   | 207 g      | 190 e  |
| 193 e   | 208 g      | 190 e  |
| 193 f   | 209 g      | 190 e  |
| 193 g   | 210 g      | 190 e  |
| 193 h   | 211 g      | 190 e  |
| 193 i   | 212 g      | 190 e  |
| 193 j   | 213 g      | 190 e  |
| 193 k   | 214 g      | 190 e  |
| 193 l   | 215 g      | 190 e  |
| 193 m   | 216 g      | 190 e  |
| 193 n   | 217 g      | 190 e  |
| 193 o   | 218 g      | 190 e  |
| 193 p   | 219 g      | 190 e  |
| 193 q   | 220 g      | 190 e  |
| 193 r   | 221 g      | 190 e  |
| 193 s   | 222 g      | 190 e  |
| 193 t   | 223 g      | 190 e  |
| 193 u   | 224 g      | 190 e  |
| 193 v   | 225 g      | 190 e  |
| 193 w   | 226 g      | 190 e  |
| 193 x   | 227 g      | 190 e  |
| 193 y   | 228 g      | 190 e  |
| 193 z   | 229 g      | 190 e  |
| 193 aa  | 230 g      | 190 e  |
| 193 ab  | 231 g      | 190 e  |
| 193 ac  | 232 g      | 190 e  |
| 193 ad  | 233 g      | 190 e  |
| 193 ae  | 234 g      | 190 e  |
| 193 af  | 235 g      | 190 e  |
| 193 ag  | 236 g      | 190 e  |
| 193 ah  | 237 g      | 190 e  |
| 193 ai  | 238 g      | 190 e  |
| 193 aj  | 239 g      | 190 e  |
| 193 ak  | 240 g      | 190 e  |
| 193 al  | 241 g      | 190 e  |
| 193 am  | 242 g      | 190 e  |
| 193 an  | 243 g      | 190 e  |
| 193 ao  | 244 g      | 190 e  |
| 193 ap  | 245 g      | 190 e  |
| 193 aqu | 246 g      | 190 e  |
| 193 ar  | 247 g      | 190 e  |
| 193 as  | 248 g      | 190 e  |
| 193 at  | 249 g      | 190 e  |
| 193 au  | 250 g      | 190 e  |
| 193 av  | 251 g      | 190 e  |
| 193 aw  | 252 g      | 190 e  |
| 193 ax  | 253 g      | 190 e  |
| 193 ay  | 254 g      | 190 e  |
| 193 az  | 255 g      | 190 e  |
| 193 ba  | 256 g      | 190 e  |
| 193 bb  | 257 g      | 190 e  |
| 193 bc  | 258 g      | 190 e  |
| 193 bd  | 259 g      | 190 e  |
| 193 be  | 260 g      | 190 e  |
| 193 bf  | 261 g      | 190 e  |
| 193 bg  | 262 g      | 190 e  |
| 193 bh  | 263 g      | 190 e  |
| 193 bi  | 264 g      | 190 e  |
| 193 bj  | 265 g      | 190 e  |
| 193 bk  | 266 g      | 190 e  |
| 193 bl  | 267 g      | 190 e  |
| 193 bm  | 268 g      | 190 e  |
| 193 bn  | 269 g      | 190 e  |
| 193 bo  | 270 g      | 190 e  |
| 193 bp  | 271 g      | 190 e  |
| 193 bq  | 272 g      | 190 e  |
| 193 br  | 273 g      | 190 e  |
| 193 bs  | 274 g      | 190 e  |
| 193 bt  | 275 g      | 190 e  |
| 193 bu  | 276 g      | 190 e  |
| 193 bv  | 277 g      | 190 e  |
| 193 bw  | 278 g      | 190 e  |
| 193 bx  | 279 g      | 190 e  |
| 193 by  | 280 g      | 190 e  |
| 193 bz  | 281 g      | 190 e  |
| 193 ca  | 282 g      | 190 e  |
| 193 cb  | 283 g      | 190 e  |
| 193 cc  | 284 g      | 190 e  |
| 193 cd  | 285 g      | 190 e  |
| 193 ce  | 286 g      | 190 e  |
| 193 cf  | 287 g      | 190 e  |
| 193 cg  | 288 g      | 190 e  |
| 193 ch  | 289 g      | 190 e  |
| 193 ci  | 290 g      | 190 e  |
| 193 cj  | 291 g      | 190 e  |
| 193 ck  | 292 g      | 190 e  |
| 193 cl  | 293 g      | 190 e  |
| 193 cm  | 294 g      | 190 e  |
| 193 cn  | 295 g      | 190 e  |
| 193 co  | 296 g      | 190 e  |
| 193 cp  | 297 g      | 190 e  |
| 193 cq  | 298 g      | 190 e  |
| 193 cr  | 299 g      | 190 e  |
| 193 cs  | 300 g      | 190 e  |
| 193 ct  | 301 g      | 190 e  |
| 193 cu  | 302 g      | 190 e  |
| 193 cv  | 303 g      | 190 e  |
| 193 cw  | 304 g      | 190 e  |
| 193 cx  | 305 g      | 190 e  |
| 193 cy  | 306 g      | 190 e  |
| 193 cz  | 307 g      | 190 e  |
| 193 da  | 308 g      | 190 e  |
| 193 db  | 309 g      | 190 e  |
| 193 dc  | 310 g      | 190 e  |
| 193 dd  | 311 g      | 190 e  |
| 193 de  | 312 g      | 190 e  |
| 193 df  | 313 g      | 190 e  |
| 193 dg  | 314 g      | 190 e  |
| 193 dh  | 315 g      | 190 e  |
| 193 di  | 316 g      | 190 e  |
| 193 dj  | 317 g      | 190 e  |
| 193 dk  | 318 g      | 190 e  |
| 193 dl  | 319 g      | 190 e  |
| 193 dm  | 320 g      | 190 e  |
| 193 dn  | 321 g      | 190 e  |
| 193 do  | 322 g      | 190 e  |
| 193 dp  | 323 g      | 190 e  |
| 193 dq  | 324 g      | 190 e  |
| 193 dr  | 325 g      | 190 e  |
| 193 ds  | 326 g      | 190 e  |

Query Match 25.3%; Score 432.4; DB 12; Length 792;  
Best Local Similarity 82.8%; Pred. No. 2e-106;  
Matches 596; Conservative 0; Mismatches 101; Indels 23; Gaps 8;

Qy 939 CTGTGTCTGCCAGACCTCCCTGGCAGCCAGGAGCCACAGCTCATGGAACAGTGTTC 998  
Db 1 CTGTGTCTGCCAGAGCTCCCTGGCATCCAGGAGCCAGCGCTCATGGAACAGTGTTC 60

Qy 999 TAAGGCCA-CTGGGATCGGATACAGAGCCCTGCCATCTCAGCAATACCCCTAGGAC 1057  
Db 61 TAGAGCCATTGTGGGATGTGAATACAGGCGCTGCCATCTCAGCAATACCCCTAGGAC 120

Qy 1058 CAGGCCCTGGACTGCTGTAGATGCTCTTCTCCATGCTCTTGGCTGAGTAACTGTGCT 1117  
Db 121 CAGGCCCTGGACTGATGATGCTCCCTCTCCATGCTCTTGGCTGAGTAACTGTGCT 180

Qy 1118 GCTTGGATCTAGTGTAGAGCACTGAGGATCCCTGCTCTGTGTAGGTAGGCGAGC 1177  
Db 181 GGTGTGATTCAGTCTCAGAGCTCTGAGCGATCCCTGCTCTGTGTAGGTAGGCGAGC 240

Qy 1178 AGTGACCACTCCAGTCTCTGG--TTACAGCTGGAATCTCTGGTCTCTGCTGCTCAT 1235  
Db 241 AGTGACCACTCAATCTCTGGATCTGACCCAGGCTCTGGGTTCTCTGCTGCTAGTT 300

Qy 1236 CCTCAAAACATCCCTGTACAAAAGGACAAACCAAGAGACCTTTAAACCTTAGGCTACT 1295  
Db 301 CCTCAAAACATCCCTGTATGAAAAGGACAAACCAAGAGACCTTTAAAGCTTAAGTACT 360

Qy 1296 GGGCAACCTGGCCAGCCTGCTGGGATAGGTAAGTCAATGTTAGG-ACCAGACAGATG 1354  
Db 361 GGGCAACCTGGG-CACATGCTGGGATAGTACAGTAATAGGTACCAGGCGAGATG 419

Qy 1355 CCTGAAATCTTCAATCTTGGACTTCTGTATGCTTGTCCCAAGATGATGAATG 1414  
Db 420 CTGGAACATCCAGGTCCTTGGACTTCTATGCTTGTCTCC-AAAATATGGGTG 478

Qy 1415 AACTCGTAAGTACCTTCCCTGACCTGAGAACACCTGCTGCTCGGGAAGTATTGAG 1474  
Db 479 ACCTTGAAGTGTGCTTCTGATCTGAGAACACCTGCGGCT----- 523

Qy 1475 GGCAGAAATCTCTGTGAACATGAAGAT-GAATCAGACTGTCTTAAAGAAATCTCGAA 1533  
Db 524 GGAAGAAATTTCTGGGAACATGAAGAGATGAATCAGACTATTCTTAAAGCGCTGGCCA 583

Qy 1534 AGTCCAGAACTTGAGCTTTGTATTTTTCAGGAATGCACATCTCTTAAGCACTCGCAAAC 1593  
Db 584 AGTCCAGAACTTGAGCTGTATGTAAACATATACATCTCTTAA-ATGCTCACAA 642

Qy 1594 AGGAAGCTCCACACCTCTGGCAGGCCAGGCGCTTTCTTTCAGCATGAGAAAGACAAG 1653  
Db 643 GCAAGAGTCCACACTTCTGGCAGGCCAGGCGCTTTCTTTCAGCATGAGAGAGACAAG 702

RESULT 8  
BI457141  
LOCUS  
DEFINITION 603185392F1 NIH\_MGC\_42 Homo sapiens cDNA clone IMAGE:5258159 5',  
mRNA sequence.  
ACCESSION BI457141  
VERSION BI457141.1 GI:15247797  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1. (bases 1 to 651)  
NIH-MGC http://mgi.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgaps-femail.nih.gov  
Tissue Procurement: ATCC

cdNA Library Preparation: Ling Hong/Rubin Laboratory  
cdNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: L1CMI869 row: g column: 24  
High quality sequence stop: 651.

FEATURES  
Location/Qualifiers  
1..651  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5258159"  
/clone\_lib="NIH\_MGC\_42"  
/tissue\_type="epithelioid carcinoma cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: pancreas; Vector: pORF7; Site\_1: XhoI;  
Site\_2: EcoRI; cDNA made by oligo-dT priming.  
Directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGCAGAG(G). Size-selected >500bp  
for average insert size 1.8kb. Library constructed by Ling  
Hong in the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).  
Note: This is a NIH\_MGC Library. |"  
BASE COUNT 91 a 230 c 219 g 111 t  
ORIGIN

Query Match 25.2%; Score 430.2; DB 13; Length 651;  
Best Local Similarity 80.1%; Pred. No. 7.3e-106;  
Matches 518; Conservative 0; Mismatches 128; Indels 1; Gaps 1;

Qy 285 TCCTCTCTCTGCTCTCTCAATGGTGTGTCGCCAGCTGTGCGGACACCTGTACTGT 344  
Db 6 TCCTCTCTCTGCTCTCTCAAGGTGCTGACCAAGTGTGCGGACACCTGTACTGT 65

Qy 345 CCTTGGACACACCCAGTGCACAGGGGTACCCCTGTGTGATGCTGTGGCTGC 404  
Db 66 CCCTGGCACCTCCCGATGCGCGCTGGGAGTACCCCTGTGTGATGCTGTGGCTGC 125

Qy 405 TGTAAAGTGTGCACGGAGCTGGGGAGTCTTGGACCACTGTGATGCTGCGGCC 464  
Db 126 TGCCTGGTATGTGCACGGGCTGGGGAGCCCTTGCACCACTCCACGCTCGGAGCC 185

Qy 465 AGCAGGCTCTGTTGTACGCTGGGAGGCTTGGGAGCCCTGCGGAGTGTGTGCTC 524  
Db 186 AGCA-GGCTGTGTGCGAGCCGCGGAGCCGCTGGAGGGGGGCTGTGCTC 244

Qy 525 TTGGATGAGGATGAGGTAGCTGTGAGTGAATGGCCGAGGTACCTGGAGGAGACC 584  
Db 245 TTGGCAGAGCAGCAGCAGCTGTAGTGAACGGCGCTGTATCGGAAGGGAGACC 304

Qy 585 TTTAAACCCAAATTCAGGGTCTTGTGCCGTGTGATGACGCTTCACTGCTGCCG 644  
Db 305 TTCCAGCCCACTGACGATCCCTGCTGGGAGAGCGGCTTCACTGCTGCCG 364

Qy 645 CTGTGAGTGAAGTGTGCGCTGCGCAGCTGCGGAGTGCACCGCCGCAAGAAATACAG 704  
Db 365 CTGTGAGCAGGATGTGCGCTGCGCAGCTGCGGAGTGCACCGCCGAGGGTGCAG 424

Qy 705 GTCCAGGAAAGTGTGCGCCGAGTGGTATGTGACAGGAGTGTGACACCGGAGTCCAG 764  
Db 425 GTCTGGGCAAGTGTGCTGCCGTGAGTGGGTGCGGCAAGGAGGGGACTGGGACCCAG 484

Qy 765 CGCTCCAGCGCGCAAGGACACCACTTCTGCGCTTGTCTACTCTGCTGCTGTGCT 824  
Db 485 CCCTTCCAGCCCAAGGACCCAGTTTCTGGCTTGTCTCTCTGCTGCTGCTGTGT 544

Qy 825 CCTTGTCCAAATTTGGAGCAGCAGCTGGGGCCCTGCTCAACACCTGCTGGCTGGGCATA 884  
Db 545 CCCTGCCCAAGTGGAGCAGCGCTGGGAGCCCTGCTCGACCACTGCTGGGCTGGCATG 604

Qy 885 GCCACCCGAGTGTCCAAACCAAGACCGATTTCTGCCAACTGGAGATCCA 931

```

Db      605  GGCACCGGGTGTCCAGCAGAACCGGGTCTCCGAGCTGGAGACCCA 651
|||||
RESULT 9
BF141695
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BF141695
601790752F1 NCI_CGAP_Lu30 Mus musculus cDNA clone IMAGE:4021308 5',
mRNA sequence.
BF141695
BF141695.1 GI:10980735
EST.
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 966)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9276 row: h column: 13
High quality sequence stop: 613.
Location/Qualifiers
1..966
/organism="Mus musculus"
/strain="CZECH II"
/db_xref="taxon:10090"
/clone="IMAGE:4021308"
/tissue_type="tumor, metastatic to mammary"
/lab_host="DH10B"
/note="Organ: Lung; Vector: pCMV-SPORT6; Site:1: NotI;
Site:2: SalI; transgenic model WNT-1, expression driven by
MTWV-LTR enhancer; Cloned unidirectionally. Primer: Oligo
dt. Library constructed by Life Technologies.
Investigator providing samples: Gilbert Smith, NIH"

BASE COUNT 227 a 298 c 258 g 183 t
ORIGIN
source

FEATURES
Query Match 25.1%; Score 429.4; DB 12; Length 966;
Best Local Similarity 85.6%; Pred. No. 1.4e-105;
Matches 571; Conservative 0; Mismatches 86; Indels 10; Gaps 8;

QY 808 CTGCTCTGCTGATGCTCTTCCAAATTGAGCAGACCGCTGGGGCCCTCTCAACCA 867
|||||
Db 1 CTGCATCTGCCATGGCCCTGTCCAACTGAGCAGACCGCTGGGGCCCTCTCAACCA 60
QY 868 CTTGTGGGCTGGGCATAGCACCCCGAGTGTCACCAAGAACCGATTCTGCCAACTGGAGA 927
|||||
Db 61 CTTGTGGGCTGGGCATAGCACCCCGAGTGTCACCAAGAACCGATTCTGCCAACTGGAGA 120
QY 928 TCCACAGCGCGCTGTGTGTCGCCACAGCCCTGGCTGGCAGCAGGACACGCTCATGGA 987
|||||
Db 121 TCCACAGCGCGCTGTGTGTCGCCACAGCCCTGGCTGGCAGCAGGACACGCTCATGGA 180
QY 988 ACAGTGCCTTCTTAAGGCCA-ACCTGGGGATGCGGATACAGGGCTCGCCATCCFCACCAAT 1046
|||||
Db 181 ACAGTGCCTTCTTAGAGCCATTCTGGGGATGTGAATACAGGGCTCGCCATTCACCAAT 240
QY 1047 GACCTTAGACACAGCCCTGGGACTGCTGTAGATGCTCTTCTCCATGCTCTTTGGCTGAC 1106
|||||
Db 241 GTCCCTTAGACACAGCCCTGGGACTGATGATGATGATGATGATGATGATGATGATGATG 300
QY 1107 TTAACGTCTCTTGGATTCACTGTGTAGACCCACTGAGCGGATCCCTGCTGCTGTCTGAG 1166
|||||
Db      301  TTAAGTGTCTCTGGGTGGGATTCAGTGTCCAGAGCCCTTGAGCGATCCCTGCTGTCTGAG 360
|||||
QY 1167 GTAGCGGAGAGGTGACCAGCTCCAGTCTCTTCTGG--TTACAGCTGGGAATTTGGTCT 1224
|||||
Db 361 GTGGGGGAGAGGTGACCAGCTACATTTCTCTGATTTGACCCAGGCTTCTGGTCT 420
|||||
QY 1225 CTTGCTCATTTCTCTCAAAACATCCCTGTACAAAAAGGACAAACAAAAAGACCTTTAAACC 1284
|||||
Db 421 CTTGCTCATTTCTCTCAAAACATCCCTGTATGAAAGGACAAACAAAAAGACCTTTAAAGC 480
|||||
QY 1285 TAGGCTATCTGGGGAACCTGGCCACCTGGGATAGGTAAGTCAATTTAGG-ACCAG 1343
|||||
Db 481 TAAGTGTACT-GGCAAGCCTGGCCACCATGCTGGGGATGACAGTAAATAGGTACCAAG 539
|||||
QY 1344 ACAGCAGATTTGCTGAAACTTCCAAATTCCTTTCTTGGAGCTTCTGTATGCTGTGCCCAAA 1403
|||||
Db 540 GCAGCAGAT--CCTGAACATCCAGTCCCTTCTTGGAGCTTCTGTATGCTGTGCCCAAA 596
|||||
QY 1404 GATGATGAATGAATCGTAAAGTGTACCTTCCCTGACCTGAGAACACCTGCTGCTCGGG 1463
|||||
Db 597 GATTATGGGTGACCT-GTAAGTGTGCC-TCCCTGATCTGAGAACACCTGGCCGGGGGC 654
|||||
QY 1464 AAGTATT 1470
|||||
Db 655 ACGAAAT 661
|||||

RESULT 10
BQ937887
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BQ937887
940 bp mRNA linear EST 21-AUG-2002
AGENCOURT_8951807 NCI_CGAP_Co24 Mus musculus cDNA clone
IMAGE:6476852 5', mRNA sequence.
BQ937887
BQ937887.1 GI:22353365
EST.
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 940)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue procurement: The Cepko Laboratory
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM14017 row: n column: 21
High quality sequence stop: 543.
Location/Qualifiers
1..940
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:6476852"
/lab_host="NCI_CGAP_Co24"
/note="Organ: colon; Vector: pCMV-SPORT6; Site:1: NotI;
Site:2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.6 kb. Constructed by Life
Technologies. Note: this is a NCI-CGAP Library."

BASE COUNT 169 a 277 c 288 g 200 t
ORIGIN
source

FEATURES
Query Match 24.9%; Score 425.8; DB 14; Length 940;
Best Local Similarity 90.9%; Pred. No. 1.3e-104;
Matches 488; Conservative 0; Mismatches 39; Indels 10; Gaps 3;

```



[illegible]







GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 15:59:01 ; Search time 486.369 Seconds  
(without alignments)  
10471.464 Million cell updates/sec

Title: US-10-010-408-1\_COPY\_1534\_1708  
Perfect score: 175  
Sequence: 1 AGTCCAGGAACCTTGAGCTTT.....GCTAGATAAACAACCCAAA 175

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:\*\*

- 1: gb\_ba:\*\*
- 2: gb\_htg:\*\*
- 3: gb\_in:\*\*
- 4: gb\_om:\*\*
- 5: gb\_ov:\*\*
- 6: gb\_pat:\*\*
- 7: gb\_ph:\*\*
- 8: gb\_pi:\*\*
- 9: gb\_pr:\*\*
- 10: gb\_ro:\*\*
- 11: gb\_sts:\*\*
- 12: gb\_sy:\*\*
- 13: gb\_un:\*\*
- 14: gb\_vi:\*\*
- 15: em\_ba:\*\*
- 16: em\_fun:\*\*
- 17: em\_hum:\*\*
- 18: em\_in:\*\*
- 19: em\_mu:\*\*
- 20: em\_om:\*\*
- 21: em\_or:\*\*
- 22: em\_ov:\*\*
- 23: em\_pat:\*\*
- 24: em\_ph:\*\*
- 25: em\_pi:\*\*
- 26: em\_ro:\*\*
- 27: em\_sts:\*\*
- 28: em\_un:\*\*
- 29: em\_vi:\*\*
- 30: em\_htg\_hum:\*\*
- 31: em\_htg\_inv:\*\*
- 32: em\_htg\_other:\*\*
- 33: em\_htg\_mus:\*\*
- 34: em\_htg\_pln:\*\*
- 35: em\_htg\_rod:\*\*
- 36: em\_htg\_mam:\*\*
- 37: em\_htg\_vrt:\*\*
- 38: em\_sy:\*\*
- 39: em\_htgo\_hum:\*\*
- 40: em\_htgo\_mus:\*\*
- 41: em\_htgo\_other:\*\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description        |
|------------|-------|-------------|--------|-------|--------------------|
| 1          | 171.8 | 98.2        | 137964 | 2     | AC126895 Rattus no |
| 2          | 171.8 | 98.2        | 226303 | 2     | AC095418 Rattus no |
| 3          | 156.6 | 89.5        | 1741   | 10    | AF259981 Rattus no |
| c          | 126.2 | 72.1        | 1734   | 6     | AR210324 Sequence  |
| 4          | 126.2 | 72.1        | 1734   | 6     | AR210325 Sequence  |
| 5          | 126.2 | 72.1        | 1734   | 6     | AF100778 Mus muscu |
| 6          | 126.2 | 72.1        | 61072  | 10    | AL731698 Mouse DNA |
| 7          | 126.2 | 72.1        | 216757 | 2     | AL669906 Mus muscu |
| 8          | 126.2 | 72.1        | 216757 | 2     | AL669906 Mus muscu |
| 9          | 121.4 | 69.4        | 1739   | 10    | AF126063 Mus muscu |
| c          | 36.4  | 20.8        | 208219 | 2     | AC130013 Rattus no |
| 10         | 35    | 20.0        | 112855 | 2     | AP003608 Oryza sat |
| 11         | 34.8  | 19.9        | 143565 | 2     | AC023195 Homo sapi |
| 12         | 34.8  | 19.9        | 190964 | 10    | AL671848 Mouse DNA |
| c          | 34.8  | 19.9        | 233269 | 2     | AC094291 Rattus no |
| 13         | 34.4  | 19.7        | 144233 | 2     | AC088103 Homo sapi |
| c          | 34.4  | 19.7        | 161900 | 9     | AC096729 Homo sapi |
| 14         | 33.6  | 19.2        | 238910 | 2     | AC122851 Mus muscu |
| 15         | 33.4  | 19.1        | 167633 | 2     | AC129096 Papio cyn |
| 16         | 33.4  | 18.9        | 171564 | 2     | AC113225 Rattus no |
| 17         | 33    | 18.9        | 173085 | 2     | AC112105 Rattus no |
| 18         | 33    | 18.9        | 173085 | 2     | AC112105 Rattus no |
| 19         | 33    | 18.9        | 181719 | 2     | AL662863 Mus muscu |
| 20         | 33    | 18.9        | 181719 | 2     | AL662863 Mus muscu |
| c          | 33    | 18.9        | 195339 | 10    | AL596181 Mouse DNA |
| 21         | 33    | 18.9        | 210487 | 9     | AL445199 Human DNA |
| 22         | 32.8  | 18.7        | 145832 | 9     | AC090736 Homo sapi |
| c          | 32.8  | 18.7        | 203217 | 9     | AC020603 Homo sapi |
| 23         | 32.8  | 18.7        | 203217 | 9     | AC020603 Homo sapi |
| 24         | 32.8  | 18.7        | 203226 | 2     | AC109624 Mus muscu |
| c          | 32.8  | 18.7        | 213438 | 2     | AL844489 Mus muscu |
| 25         | 32.4  | 18.5        | 65932  | 2     | AC107972 Homo sapi |
| 26         | 32.4  | 18.5        | 160199 | 9     | AF391285 Homo sapi |
| 27         | 32.4  | 18.5        | 182522 | 2     | AC015689 Homo sapi |
| 28         | 32.2  | 18.4        | 119531 | 2     | AP003879 Oryza sat |
| c          | 32.2  | 18.4        | 150428 | 2     | AP004591 Oryza sat |
| 29         | 32    | 18.3        | 173755 | 2     | AC121976 Mus muscu |
| 30         | 32    | 18.3        | 185536 | 2     | AC122340 Mus muscu |
| 31         | 32    | 18.3        | 187490 | 2     | AC048378 Homo sapi |
| 32         | 32    | 18.3        | 192404 | 2     | AC128523 Rattus no |
| 33         | 32    | 18.3        | 202634 | 2     | AC114771 Homo sapi |
| 34         | 32    | 18.3        | 212421 | 2     | AC109591 Homo sapi |
| c          | 32    | 18.3        | 221647 | 10    | AL591003 Mouse DNA |
| 35         | 31.8  | 18.2        | 227103 | 2     | AC122470 Mus muscu |
| 36         | 31.8  | 18.2        | 110000 | 2     | Continuation (3 of |
| 37         | 31.8  | 18.2        | 132278 | 9     | AC026784 Homo sapi |
| 38         | 31.8  | 18.2        | 147817 | 8     | AP003267 Oryza sat |
| 39         | 31.8  | 18.2        | 151269 | 8     | AC016781 Genomic S |
| 40         | 31.8  | 18.2        | 151491 | 2     | AC084764 Oryza sat |
| 41         | 31.8  | 18.2        | 151491 | 2     | AC084764 Oryza sat |

ALIGNMENTS

RESULT 1  
AC126895  
LOCUS  
DEFINITION  
Rattus norvegicus clone CH230-301E4, \*\*\* SEQUENCING IN PROGRESS  
\*\*\*, 49 unordered pieces.  
AC126895 137964 bp DNA linear HTG 24-JUL-2002  
AC126895  
AC126895.1 GI:21724040  
HTG; HTGS\_PHASE1.  
Rattus norvegicus  
Rattus norvegicus  
ORGANISM  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
REFERENCE 1 (bases 1 to 137964)  
AUTHORS Muzny,D.M., Adams,C., Ali-Osman,F.R., Allen,C.,

Alsbrooks, S.L., Amaraturunge, H.C., Are, J.R., Ayele, M., Banks, T.,  
 Barbaria, J., Benton, J., Bimage, K., Blankenburg, K., Bonnin, D.,  
 Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P.,  
 Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,  
 Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,  
 Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,  
 Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,  
 Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,  
 Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,  
 Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,  
 Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,  
 Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,  
 Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,  
 Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,  
 Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,  
 Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B.,  
 Homisi, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,  
 Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,  
 Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,  
 Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,  
 Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulseged, H.,  
 Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,  
 Maheshwari, M., Mapa, P., Martin, R., Martindale, A., Martinez, E.,  
 Massey, E., Mawhinney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M.,  
 Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,  
 Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,  
 Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogih, M., Okwuonu, G.,  
 Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,  
 Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,  
 Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G.,  
 Scherer, S., Scott, G., Shen, H., Shooshtari, N., Sisson, I.,  
 Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H.,  
 Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,  
 Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,  
 Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q.,  
 Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,  
 Williams, G., Williamson, A., Wleczky, R., Wooden, S., Worley, K.,  
 Wu, C., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,  
 Weinstein, G., and Gibbs, R.  
 Direct Submission  
 Unpublished  
 2 (bases 1 to 137964)  
 Worley, K.C.  
 Direct Submission  
 Submitted (10-JUL-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 137964)  
 Worley, K.C.  
 Direct Submission  
 Submitted (24-JUL-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 ----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
 ----- Project Information  
 Center project name: G2HG  
 Center clone name: CH230-301E4  
 ----- Summary Statistics  
 Sequencing vector: Plasmid;  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.990329  
 Consensus quality: 89474 bases at least Q40  
 Consensus quality: 93422 bases at least Q30  
 Consensus quality: 96506 bases at least Q20  
 -----  
 \* NOTE: Estimated insert size may differ from sequence length  
 \* (see <http://www.hgsc.bcm.tmc.edu/docs/Genbank.draft.data.html>).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 49 contigs. The true order of the pieces

\* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 \* 1  
 \* 1056: contig of 1056 bp in length  
 \* 1057 1156: gap of unknown length  
 \* 1157 2335: contig of 1179 bp in length  
 \* 2336 2435: gap of unknown length  
 \* 2436 3442: contig of 1007 bp in length  
 \* 3443 3542: gap of unknown length  
 \* 3543 5081: contig of 1539 bp in length  
 \* 5082 5181: gap of unknown length  
 \* 5182 6424: contig of 1243 bp in length  
 \* 6425 6524: gap of unknown length  
 \* 6525 7814: contig of 1290 bp in length  
 \* 7815 7914: gap of unknown length  
 \* 7915 9153: contig of 1239 bp in length  
 \* 9154 9253: gap of unknown length  
 \* 9254 10445: contig of 1192 bp in length  
 \* 10446 10545: gap of unknown length  
 \* 10546 11771: contig of 1226 bp in length  
 \* 11772 11871: gap of unknown length  
 \* 11872 13240: contig of 1369 bp in length  
 \* 13241 13340: gap of unknown length  
 \* 13341 14359: contig of 1019 bp in length  
 \* 14360 14459: gap of unknown length  
 \* 14460 15551: contig of 1192 bp in length  
 \* 15552 15751: gap of unknown length  
 \* 15752 17494: contig of 1743 bp in length  
 \* 17495 18594: gap of unknown length  
 \* 18595 18679: contig of 1085 bp in length  
 \* 18680 18779: gap of unknown length  
 \* 18780 20681: contig of 1902 bp in length  
 \* 20682 20781: gap of unknown length  
 \* 20782 22118: contig of 1337 bp in length  
 \* 22119 22218: gap of unknown length  
 \* 22219 23578: contig of 1360 bp in length  
 \* 23579 23678: gap of unknown length  
 \* 23679 25423: contig of 1745 bp in length  
 \* 25424 25523: gap of unknown length  
 \* 25524 27808: contig of 2285 bp in length  
 \* 27809 27908: gap of unknown length  
 \* 27909 30372: contig of 2364 bp in length  
 \* 30373 30372: gap of unknown length  
 \* 30373 33091: contig of 2719 bp in length  
 \* 33092 33191: gap of unknown length  
 \* 33192 35777: contig of 2586 bp in length  
 \* 35778 35877: gap of unknown length  
 \* 35878 36908: contig of 1031 bp in length  
 \* 36909 37008: gap of unknown length  
 \* 37009 38530: contig of 1922 bp in length  
 \* 38531 39030: gap of unknown length  
 \* 39031 41210: contig of 2180 bp in length  
 \* 41211 41310: gap of unknown length  
 \* 41311 42937: contig of 1627 bp in length  
 \* 42938 43037: gap of unknown length  
 \* 43038 46120: contig of 3083 bp in length  
 \* 46121 46220: gap of unknown length  
 \* 46221 48217: contig of 1997 bp in length  
 \* 48218 48317: gap of unknown length  
 \* 48318 50433: contig of 2116 bp in length  
 \* 50434 50533: gap of unknown length  
 \* 50534 52880: contig of 2347 bp in length  
 \* 52881 52980: gap of unknown length  
 \* 52981 55194: contig of 2214 bp in length  
 \* 55195 57482: gap of unknown length  
 \* 57483 57582: gap of unknown length  
 \* 57583 60781: contig of 3199 bp in length  
 \* 60782 60881: gap of unknown length  
 \* 60882 62599: contig of 1718 bp in length

```

* 62600 62699: gap of unknown length
* 62700 65203: contig of 2504 bp in length
* 62704 65303: gap of unknown length
* 63004 66914: contig of 1611 bp in length
* 63015 67014: gap of unknown length
* 67015 70839: contig of 3825 bp in length
* 70840 70939: gap of unknown length
* 70940 75139: contig of 4200 bp in length
* 75140 75239: gap of unknown length
* 75240 78296: contig of 3057 bp in length
* 78297 83138: gap of unknown length
* 83139 83238: contig of 4742 bp in length
* 83239 88204: gap of unknown length
* 88205 92238: contig of 4966 bp in length
* 92239 92338: gap of unknown length
* 92340 97339: contig of 3934 bp in length
* 97340 97439: gap of unknown length
* 97440 103535: contig of 5001 bp in length
* 103536 103634: gap of unknown length
* 103635 107080: contig of 6095 bp in length
* 107081 107180: gap of unknown length
* 107181 115201: contig of 3446 bp in length
* 115202 115301: gap of unknown length
* 115302 121237: contig of 8021 bp in length
* 121238 121337: gap of unknown length
* 121338 127910: contig of 5936 bp in length
* 127911 128010: gap of unknown length
* 128011 137964: contig of 6573 bp in length
* 128012 137964: contig of 9954 bp in length.
FEATURES             Location/Qualifiers
     source            1..137964
                        /organism="Rattus norvegicus"
                        /db_xref="taxon:10116"
                        /clone="CH230-301E4"
BASE COUNT   31831 a 34027 c 35229 g 31979 t 4898 others
ORIGIN

Query Match      98.2%; Score 171.8; DB 2; Length 137964;
Best Local Similarity 98.9%; Pred. No. 6.9e-50;
Matches 173; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGTCCAGGACTTGACCTTTGTTATTTTCAGGAATGCACATCTCTTAAGCACTCGCAAAAC 60
Db 44819 AGTCCAGGAACTTGACCTTTGTTATTTTCAGGAATGCACATCTCTTAAGCACTCGCAAAAC
|||||
QY 61 AGGAAGGCTCCACACCTCTCGCAGGCGGCGCTTCTCTTCAGCATGAGAAAGACAAGG 120
Db 44879 AGGAAGGCTCCACACCTCTCGTAGGCCAGGCGCTTCTCTTCAGCATGAGAAAGACAAGG 44938
|||||
QY 121 GACAGCAGACTACTCTCTCTGAGGACTAGTCTAGCCTAGATAACACCCCAA 175
Db 44939 GACAGCAGACTACTCTCTCTGAGGACTAGTCTAGCCTAGATAACACCCCAA 44993
|||||

RESULT 2
AC095418
LOCUS
DEFINITION
Rattus norvegicus clone CH230-7c10, *** SEQUENCING IN PROGRESS ***,
51 unordered pieces.
AC095418
VERSION
AC095418.3 GI:21717893
KEYWORDS
HTG; HTGS_PHASE1.
SOURCE
Norway rat.
ORGANISM
Rattus norvegicus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 226303)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amarantunge,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbaria,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D.,
Bourca,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,

```

```

Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
DeLaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlssohn,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lied,C., Liu,J., Liu,W., Loulseghe,H.,
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,B., Nwokenkwo,S., Ogih,M., Okwona,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savory,G.,
Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,
Sodergren,E., Sonaik,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Teifrod,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 226303)
Worley,K.C.
Direct Submission
Submitted (17-SEP-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 226303)
Worley,K.C.
Direct Submission
Submitted (11-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 10, 2002 this sequence version replaced gi:17941885.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GCDF
Center clone name: CH230-7C10
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 185950 bases at least Q40
Consensus quality: 190362 bases at least Q30
Consensus quality: 193076 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank.draft.data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 51 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as

```

\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 1051: contig of 1051 bp in length  
1052 1151: gap of unknown length  
1152 2286: contig of 1135 bp in length  
2287 2386: gap of unknown length  
2387 3433: contig of 1047 bp in length  
3434 3533: gap of unknown length  
3534 4614: contig of 1081 bp in length  
4615 4714: gap of unknown length  
4715 6581: contig of 1867 bp in length  
6582 6681: gap of unknown length  
6682 7773: contig of 1092 bp in length  
7774 7873: gap of unknown length  
7874 9530: contig of 1657 bp in length  
9531 9630: gap of unknown length  
9632 10661: contig of 1031 bp in length  
10662 10761: gap of unknown length  
10762 11958: contig of 1197 bp in length  
11959 12058: gap of unknown length  
12059 13140: contig of 1082 bp in length  
13141 13240: gap of unknown length  
13241 14777: contig of 1537 bp in length  
14778 14877: gap of unknown length  
14879 16063: contig of 1186 bp in length  
16064 16163: gap of unknown length  
16164 17585: contig of 1422 bp in length  
17586 17685: gap of unknown length  
17686 18916: contig of 1231 bp in length  
18917 19016: gap of unknown length  
19017 20305: contig of 1289 bp in length  
20306 20405: gap of unknown length  
20406 21537: contig of 1132 bp in length  
21538 21637: gap of unknown length  
21638 23703: contig of 2066 bp in length  
23704 23803: gap of unknown length  
23804 26306: contig of 2503 bp in length  
26307 26406: gap of unknown length  
26407 28431: contig of 2025 bp in length  
28432 28531: gap of unknown length  
28532 30615: contig of 2084 bp in length  
30616 30715: gap of unknown length  
30716 33215: contig of 2500 bp in length  
33216 33315: gap of unknown length  
33316 35357: contig of 2042 bp in length  
35358 35457: gap of unknown length  
35458 38012: contig of 2355 bp in length  
38013 38112: gap of unknown length  
38113 40338: contig of 2226 bp in length  
40339 40438: gap of unknown length  
40439 44051: contig of 3613 bp in length  
44052 44151: gap of unknown length  
44152 46294: contig of 2143 bp in length  
46295 46394: gap of unknown length  
46395 48845: contig of 2451 bp in length  
48846 48945: gap of unknown length  
48946 51753: contig of 2808 bp in length  
51754 51853: gap of unknown length  
51854 55915: contig of 4062 bp in length  
55916 56015: gap of unknown length  
56016 60468: contig of 4453 bp in length  
60469 60568: gap of unknown length  
60569 63640: contig of 3072 bp in length  
63641 63740: gap of unknown length  
63741 67888: contig of 4148 bp in length  
67889 67988: gap of unknown length  
67989 73014: contig of 5026 bp in length  
73015 73114: gap of unknown length  
73115 77313: contig of 4199 bp in length  
77314 77414: gap of unknown length  
83786: contig of 6373 bp in length

\* 83787 83886: gap of unknown length  
83887 85990: contig of 5704 bp in length  
85991 89690: gap of unknown length  
89691 93846: contig of 4156 bp in length  
93847 93946: gap of unknown length  
93947 98126: contig of 4180 bp in length  
98127 98226: gap of unknown length  
98227 104258: contig of 6032 bp in length  
104259 104358: gap of unknown length  
104359 110994: contig of 6636 bp in length  
110995 111094: gap of unknown length  
111096 119838: contig of 8744 bp in length  
119839 126339: gap of unknown length  
126340 126439: gap of unknown length  
126440 134888: contig of 8449 bp in length  
134889 134988: gap of unknown length  
134989 142709: contig of 7721 bp in length  
142710 142809: gap of unknown length  
142810 149818: contig of 7009 bp in length  
149819 149818: gap of unknown length  
149819 158719: contig of 8801 bp in length  
158720 158819: gap of unknown length  
158820 169868: contig of 11049 bp in length  
169869 169968: gap of unknown length  
169969 180390: contig of 10422 bp in length  
180391 180490: gap of unknown length  
180491 191272: contig of 10782 bp in length  
191273 191372: gap of unknown length  
191373 209667: contig of 18295 bp in length  
209668 209767: gap of unknown length  
209768 226303: contig of 16536 bp in length.

FEATURES  
Location/Qualifiers  
1..226303  
/organism="Rattus norvegicus"  
/db\_xref="taxon:10116"  
/clone="CH230-7C10"

BASE COUNT 57875 a 53353 c 52917 g 57041 t 5117 others

Query Match 98.2%; Score 171.8; DB 2: Length 226303;  
Best Local Similarity 98.9%; Pred No. 7,5e-50;  
Matches 173; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AGTCCAGGAACCTGAGCTTTTGTATTTTCAGGAATGCACATCTCTTAAGCACTCGCAAAAC 60  
|||||  
Db 113012 AGTCCAGGAACCTGAGCTTTTGTATTTTCAGGAATGCACATCTCTTAAGCACTCGCAAAAC 113071  
|||||

Qy 61 AGGAAGGCTCCACACCTCTGCGAGGCCAGGCCCTTCTCTTCAGCATGAGAAGACAGG 120  
|||||

Db 113072 AGGAAGGCTCCACACCTCTGCGAGGCCAGGCCCTTCTCTTCAGCATGAGAAGACAGG 113131  
|||||

Qy 121 GACAGCAGAGTACTCTCTCTGGAGGACTAGTCTAGCCCTAGATTAATACACCCAAA 175  
|||||

Db 113132 GACAGCAGAGTACTCTCTCTGGAGGACTAGTCTAGCCCTAGATTAATACACCCAAA 113186  
|||||

## RESULT 3

AF259981

LOCUS

DEFINITION

AF259981.1

ACCESSION

AF259981.1

KEYWORDS

SOURCE

ORGANISM

Rattus norvegicus.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

REFERENCE

AUTHORS

TITLE

Identification of rCop-1, a new member of the CCN protein family,

AF259981 1741 bp mRNA linear ROD 09-MAY-2000  
Rattus norvegicus CCN family protein COP-1 (Cop-1) mRNA, complete cds.

AF259981.1 GI:7739780  
Rattus norvegicus.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
1 (bases 1 to 1741)  
Zhang, R., Averbouch, L., Zhu, W., Zhang, H., Jo, H., Dempsey, P.J.,  
Coffey, R.J., Pardee, A.B., and Liang, P.  
Identification of rCop-1, a new member of the CCN protein family,

as a negative regulator for cell transformation  
Mol. Cell. Biol. 18 (10), 6131-6141 (1998)

JOURNAL  
MEDLINE

98414629

REFERENCE

2 (bases 1 to 1741)

Liang, P.

Direct Submission

Submitted (24-APR-2000) Cell Biology, Vanderbilt-Ingram Cancer  
Center, 649 MRB II, Nashville, TN 37232, USA

JOURNAL

TITLE

Location/Qualifiers

1..1741

source

gene

CDS

FEATURES

source

1..1741

ORGANISM

Unknown.

REFERENCE

1 (bases 1 to 1741)

AUTHORS

Botstein, D.A., Cohen, R.L., Goddard, A.D., Gurney, A.L., Hillan, K.J.,

Lawrence, D.A., Levine, A.J., Pennica, D., Roy, M. Ann. and Wood, W.I.

WSP polypeptides and nucleic acids encoding same

Patent: US 6387657-A 17 14-MAY-2002;

Location/Qualifiers

1..1734

source

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

Matches 170; Conservative

0; Mismatches 4; Indels 1; Gaps 1;

Qy 1 AGTCAGGAACTTGGAGCTTTGATTTTTCAGGAATGCACATCTCTTAAGCACTCGCAAAAC 60

Db 1547 AGTCAGGAACTTGGAGCTTTGATTTTTCAGGAATGCACATCTCTTAAGCACTCGCAAAAC 1606

Qy 61 AGGAAGGCTCCACACCTCTTAACA-GCCAGGCGCTTCTCTTCAGCATGAGAAACACAGG 120

Db 1607 AGGAAGGCTCCACACCTCTTAACA-GCCAGGCGCTTCTCTTCAGCATGAGAAACACAGG 1665

Qy 121 GACAGCAGAGTACTCTCTCTGGAGGACTAGTCTAGCTAGATAGATAAACACCCAAA 175

Db 1666 GACAGCAGAGTACTCTCTCTGGAGGACTAGTCTAGCTAGATAGATAAACACCCAAA 1720

RESULT 4

AR210324

LOCUS

Sequence 17 from patent US 6387657.

Accession AR210324

Version AR210324.1 GI:21512525

Keywords

Unknown.

ORGANISM

Unknown.

REFERENCE

1 (bases 1 to 1734)

AUTHORS

Botstein, D.A., Cohen, R.L., Goddard, A.D., Gurney, A.L., Hillan, K.J.,

Lawrence, D.A., Levine, A.J., Pennica, D., Roy, M. Ann. and Wood, W.I.

WSP polypeptides and nucleic acids encoding same

Patent: US 6387657-A 17 14-MAY-2002;

Location/Qualifiers

1..1734

source

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

Matches 151; Conservative

0; Mismatches 23; Indels 1; Gaps 1;

Qy 1 AGTCAGGAACTTGGAGCTTTGATTTTTCAGGAATGCACATCTCTTAAGCACTCGCAAAAC 60

1536 AGTCAGGAACTTGGAGCTTTGATTTTTCAGGAATGCACATCTCTTAATGCTCAAAAGC 1595

Qy 61 AGGAAGGCTCCACACCTCTTAACA-GCCAGGCGCTTCTCTTCAGCATGAGAAACACAGG 120

Db 1596 AAG-AGGCTCCACACTTCTGGAGGCGCTTCTCTTCAGCATGAGAGACAGG 1654

Qy 121 GACAGCAGAGTACTCTCTCTGGAGGACTAGTCTAGCTAGATAGATAAACACCCAAA 175

Db 1655 AACAGTAGAGTACCCTCTCTGGAGGACTGCCCGGTCTGGAATAAACACCCAAA 1709

RESULT 5

AR210325/c

LOCUS

Sequence 18 from patent US 6387657.

Accession AR210325

Version AR210325.1 GI:21512526

Keywords

Unknown.

ORGANISM

Unknown.

REFERENCE

1 (bases 1 to 1734)

AUTHORS

Botstein, D.A., Cohen, R.L., Goddard, A.D., Gurney, A.L., Hillan, K.J.,

Lawrence, D.A., Levine, A.J., Pennica, D., Roy, M. Ann. and Wood, W.I.

WSP polypeptides and nucleic acids encoding same

Patent: US 6387657-A 18 14-MAY-2002;

Location/Qualifiers

1..1734

source

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

Matches 151; Conservative

0; Mismatches 23; Indels 1; Gaps 1;

Qy 1 AGTCAGGAACTTGGAGCTTTGATTTTTCAGGAATGCACATCTCTTAATGCTCAAAAC 60

Db 199 AGTCAGGAACTTGGAGCTTTGATTTTTCAGGAATGCACATCTCTTAATGCTCAAAAC 140

Qy 61 AGGAAGGCTCCACACCTCTTAACA-GCCAGGCGCTTCTCTTCAGCATGAGAAACACAGG 120

Db 139 AAG-AGGCTCCACACTTCTGGAGGCGCTTCTCTTCAGCATGAGAGACAGG 81

Qy 121 GACAGCAGAGTACTCTCTCTGGAGGACTAGTCTAGCTAGATAGATAAACACCCAAA 175

Db 80 AACAGTAGAGTACCCTCTCTGGAGGACTGCCCGGTCTGGAATAAACACCCAAA 26

RESULT 6

AF100778

LOCUS

Mus musculus connective tissue growth factor related protein WISP-2

(Wisp2) mRNA, complete cds.

Accession AF100778

Version AF100778.1 GI:4028578

Keywords

Mus musculus.

ORGANISM

Mus musculus

REFERENCE

1 (bases 1 to 1734)

AUTHORS

Pennica, D., Swanson, T.A., Welsh, J.W., Roy, M.A., Lawrence, D.A.,

Lee, J., Brush, J., Taneyhill, L.A., Deuel, B., Lew, M., Watanabe, C.,

Cohen, R.L., Melhem, M.F., Finley, G.G., Quirke, P., Goddard, A.D.,

Hillan, K.J., Gurney, A.L., Botstein, D. and Levine, A.J.

WSP genes are members of the connective tissue growth factor

family that are up-regulated in wnt-1-transformed cells and

aberrantly expressed in human colon tumors

Proc. Natl. Acad. Sci. U.S.A. 95 (25), 14717-14722 (1998)

99061933

PUBMED

9843955

REFERENCE 2 (bases 1 to 1734)  
 AUTHORS Pennica,D.  
 TITLE Direct Submission  
 JOURNAL Submitted (23-OCT-1998) Molecular Oncology, Genentech, Inc., 1 DNA Way, South San Francisco, CA 94080, USA

FEATURES  
 source Location/Qualifiers  
 1..1734  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /cell\_line="C57MG"  
 /cell\_type="epithelial"  
 /tissue\_type="mammary"  
 /note="transformed by wnt-1"  
 1..1734  
 /gene="Wisp2"  
 257..1012  
 /gene="Wisp2"  
 /codon\_start=1  
 /product="connective tissue growth factor related protein WISP-2"  
 /protein\_id="AAC96320.1"  
 /db\_xref="GI:4028579"  
 /translation="MRGNPLHLAISFLCILSMVYSOLCPAPCAPWTPPCPGVP  
 LVLDGCGCRVCARRLGSCHLHVCDFSGLYCOPGSPGRGAVCLFEEDDGSCEV  
 NGRRLDGETFAPNCVLCRCDDGGFTCLPICSEVDRLPSPDCPRRIQVPRCCPE  
 WYCDQAVQPAIOPSSAQGHLSALVTASADGCPNWSWTAWGPCSTTCGLGIATRV  
 NONRFQLEIQRLLSLRPLASRSHGWSNAF"  
 355 a 491 c 495 g 393 t

BASE COUNT 355 a 491 c 495 g 393 t

ORIGIN

Query Match 72.1%; Score 126.2; DB 10; Length 1734;  
 Best Local Similarity 86.3%; Pred. No. 7e-34;  
 Matches 151; Conservative 0; Mismatches 23; Indels 1; Gaps 1;

QY 1 AGTCAGGAACTTGAGCTTTGATTTCAGGAATGCACATCTCTTAAGCACTCGCAAAAC 60  
 |||||  
 DB 1536 AGTCAGGAACTTGAGCTTTGATTTCAGGAATGCACATCTCTTAAGCACTCGCAAAAC 60  
 |||||

QY 61 AGGAAGGCTCCACACCTCTGGCAGGCCAGGGCTTTCTCTCAGCATGAGAAGACAAG 120  
 |||||  
 DB 1596 AAG-AGGCTCCACACTCTGGCAGGCCAGGGCTTTCTCTCAGCATGAGAAGACAAG 120  
 |||||

QY 121 GACAGCAGAGTACTCTCTCTGGAGAGTAGTCTAGCCCTAGATAAACACCCAAA 175  
 |||||  
 DB 1655 AACAGTAGAGTACCCTCTCTGGAGAGTAGTCTAGCCCTAGATAAACACCCAAA 1709  
 |||||

RESULT 7  
 AL731698  
 LOCUS Mouse DNA sequence from clone RP23-161B3 on chromosome 2, complete sequence.  
 DEFINITION  
 ACCESSION AL731698.10 GI:21214309  
 VERSION  
 KEYWORDS HTG.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 61072)  
 AUTHORS Wallis,J.  
 TITLE Direct Submission  
 JOURNAL Submitted (24-MAY-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk  
 On May 25, 2002 this sequence version replaced gi:21213601.  
 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.  
 This sequence was finished as follows unless otherwise noted: all

regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone, and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Swi., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at  
 http://www.sanger.ac.uk/projects/C-elegans/wormpep RP23-161B3 is from the RPC1-23 Mouse PAC Library  
 constructed by the group of Pieter de Jong.  
 For further details see http://www.chori.org/bacpac/home.htm  
 VECTOR: pBACE3.6.

FEATURES  
 source Location/Qualifiers  
 1..61072  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /chromosome="2"  
 /clone\_lib="RPC1-23"  
 /clone="RP23-161B3"  
 15405 a 15368 c 15033 g 15266 t

BASE COUNT 15405 a 15368 c 15033 g 15266 t

ORIGIN

Query Match 72.1%; Score 126.2; DB 10; Length 61072;  
 Best Local Similarity 86.3%; Pred. No. 1.3e-33;  
 Matches 151; Conservative 0; Mismatches 23; Indels 1; Gaps 1;

QY 1 AGTCAGGAACTTGAGCTTTGATTTCAGGAATGCACATCTCTTAAGCACTCGCAAAAC 60  
 |||||  
 DB 17748 AGTCAGGAACTTGAGCTTTGATTTCAGGAATGCACATCTCTTAAGCACTCGCAAAAC 17807  
 |||||

QY 61 AGGAAGGCTCCACACCTCTGGCAGGCCAGGGCTTTCTCTCAGCATGAGAAGACAAG 120  
 |||||  
 DB 17808 AAG-AGGCTCCACACTCTGGCAGGCCAGGGCTTTCTCTCAGCATGAGAAGACAAG 120  
 |||||

QY 121 GACAGCAGAGTACTCTCTCTGGAGAGTAGTCTAGCCCTAGATAAACACCCAAA 175  
 |||||  
 DB 17867 AACAGTAGAGTACCCTCTCTGGAGAGTAGTCTAGCCCTAGATAAACACCCAAA 17921  
 |||||

RESULT 8  
 AL669906  
 LOCUS Mus musculus chromosome 2 clone RP23-217C2, \*\*\* SEQUENCING IN DEFINITION  
 ACCESSION AL669906.5 GI:21955520  
 VERSION  
 KEYWORDS HTG; HTGS\_PHASE1; HTGS\_CANCELLED.  
 SOURCE Mus musculus.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 216757)  
 AUTHORS Sims,S.  
 TITLE Direct Submission  
 JOURNAL Submitted (23-JUL-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk  
 On Jul 25, 2002 this sequence version replaced gi:18181793.  
 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.  
 This sequence was finished as follows unless otherwise noted: all

Consensus quality: 214139 bases at least Q20  
Insert size: 214757; sum-of-contigs  
Insert size: 234243; 1.8% error; agarose-fp  
Quality coverage: 6.29x in Q20 bases; sum-of-contigs Quality  
coverage: 5.90x in Q20 bases; agarose-fp

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 21 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of 'N', but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence.  
\* as soon as it is available and the accession number will  
\* be preserved.

|         |                                      |
|---------|--------------------------------------|
| 4000:   | contig of 4000 bp in length          |
| 4100:   | gap of 100 bp                        |
| 4101:   | 15267: contig of 11167 bp in length  |
| 4102:   | 15367: gap of 100 bp                 |
| 15268:  | 34970: contig of 19603 bp in length  |
| 15368:  | 34970: contig of 19603 bp in length  |
| 34971:  | 35070: gap of 100 bp                 |
| 35071:  | 39671: contig of 4601 bp in length   |
| 39672:  | 39771: gap of 100 bp                 |
| 39772:  | 46360: contig of 6589 bp in length   |
| 46361:  | 46460: gap of 100 bp                 |
| 46461:  | 49809: contig of 3349 bp in length   |
| 49810:  | 49909: gap of 100 bp                 |
| 49910:  | 54509: contig of 4600 bp in length   |
| 54510:  | 54609: gap of 100 bp                 |
| 54610:  | 65989: contig of 11380 bp in length  |
| 65990:  | 66089: gap of 100 bp                 |
| 66090:  | 84635: contig of 18546 bp in length  |
| 84636:  | 84735: gap of 100 bp                 |
| 84736:  | 87399: contig of 2664 bp in length   |
| 87400:  | 87499: gap of 100 bp                 |
| 87500:  | 98601: contig of 11102 bp in length  |
| 98602:  | 98701: gap of 100 bp                 |
| 98702:  | 105532: contig of 6821 bp in length  |
| 105523: | 105622: gap of 100 bp                |
| 105623: | 119773: contig of 14151 bp in length |
| 119774: | 119873: gap of 100 bp                |
| 119874: | 134552: contig of 14679 bp in length |
| 134553: | 134652: gap of 100 bp                |
| 134653: | 141203: contig of 6550 bp in length  |
| 141203: | 141302: gap of 100 bp                |
| 141303: | 145808: contig of 4506 bp in length  |
| 145809: | 145908: gap of 100 bp                |
| 145909: | 151572: contig of 5664 bp in length  |
| 151573: | 151672: gap of 100 bp                |
| 151673: | 169686: contig of 18014 bp in length |
| 169687: | 169786: gap of 100 bp                |
| 169787: | 200095: contig of 30309 bp in length |
| 200096: | 200195: gap of 100 bp                |
| 200196: | 213877: contig of 13682 bp in length |
| 213878: | 213977: gap of 100 bp                |
| 213978: | 216757: contig of 2780 bp in length  |

## FEATURES

```

location/qualifiers
1. .216757
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="2"
/clone="RP23-217C2"
/clone_lib="RPCI-23"
1. .4000

```

misc\_feature

```

L: 4000
/note="assembly_fragment:03429
clone_end:SP6

```

**Misc feature**

```
4101. .13267
/note="assembly_fragment:00814
fragment chain:1"
```

misc\_feature

```
13308.  .34970
/note="assembly_fragment:00812
fragment chain:1"
```

misc\_feature

|            |   |
|------------|---|
|            | /note="assembly_fragment:03544<br>fragment_chain:1"<br>39772. .46360<br>/note="assembly_fragment:01438<br>fragment_chain:1"<br>46461. .49809<br>/note="assembly_fragment:03410<br>fragment_chain:2"<br>49910. .54509<br>/note="assembly_fragment:00811<br>fragment_chain:2"<br>54610. .65989<br>/note="assembly_fragment:02217<br>fragment_chain:2"<br>66090. .84635<br>/note="assembly_fragment:03529<br>fragment_chain:2"<br>84736. .87399<br>/note="assembly_fragment:01019<br>fragment_chain:3"<br>87500. .98601<br>/note="assembly_fragment:01170<br>fragment_chain:3"<br>98702. .105322<br>/note="assembly_fragment:02829<br>fragment_chain:3"<br>105623. .119773<br>/note="assembly_fragment:01476<br>fragment_chain:4"<br>119874. .134552<br>/note="assembly_fragment:00505<br>fragment_chain:4"<br>134653. .141202<br>/note="assembly_fragment:01951<br>fragment_chain:4"<br>141303. .145808<br>/note="assembly_fragment:02637<br>fragment_chain:5"<br>145909. .151572<br>/note="assembly_fragment:02352<br>fragment_chain:5"<br>151673. .169686<br>/note="assembly_fragment:02207<br>fragment_chain:5"<br>169787. .200095<br>/note="assembly_fragment:02849<br>fragment_chain:6"<br>200196. .213877<br>/note="assembly_fragment:01052<br>fragment_chain:6"<br>213978. .216757<br>/note="assembly_fragment:01318<br>clone_end:r7<br>vector_side:right" |
| BASE COUNT | 56425 a 51573 c 51904 g 54852 t   |
| ORIGIN     |   |

Query Match 72.1%; Score 126.2; DB 2; Length 216757;  
Best Local Similarity 86.3%; Pred. No. 1.7e-33;  
Matches 151: Conservative 0; Mismatches 23; Indels 1;

|    |       |   |       |
|----|-------|---|-------|
| Qy | 1     | AGTCCAGGAACCTTGAGCTTTGTATNTTTCAGGAATGCACATCTCTTTAAGCACTCGCAAAAC | 60    |
| Db | 88854 | AGTCCAGGAACCTTGACCTTTGTATTTGTAATAATACACATCTCTTAATGCTCACAAAGC    | 88913 |
| Qy | 61    | AGGAAGCTCCACACCTCTGGCAGCGCCGCTTCTCTCAGCATGAGAAAGACAAG           | 120   |
| Db | 88914 | AAG-AGGCTCCACACTTCTGGCAGCGCGGCTTCTCTCAGCATGAGAGAGACAAG          | 88972 |
| Qy | 121   | GACAGCAGAGTACTCTCTCTGGAGACACTAGCTAGCCTAGAAATAACACCCAAA          | 175   |
| Db | 88973 | AACAGTAGATACCTCTCTCTGGAGACCTGGCCGCTCTGGAATTAACACCCAAA           | 89027 |

```

RESULT 9
AF126063
LOCUS
DEFINITION
Mus musculus connective tissue growth factor-like protein precursor
(Ctgfl) mRNA, complete cds.
ACCESSION
AF126063
VERSION
AF126063.1 GI:4337059
SOURCE
Mus musculus.
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1739)
Kumar, S., Hand, A.T., Connor, J.R., Dodds, R.A., Ryan, P.J.,
Trill, J.J., Fisher, S.M., Nuttall, M.E., Lipshutz, D.B., Zou, C.,
Hwang, S.M., Votta, B.J., James, I.E., Rieman, D.J., Gowen, M. and
Lee, J.C.
Identification and cloning of a connective tissue growth
factor-like cDNA from human osteoblasts encoding a novel regulator
of osteoblast functions
J. Biol. Chem. 274 (24), 17123-17131 (1999)
99287915
10358067
2 (bases 1 to 1739)
Kumar, S. and Zou, C.
Direct Submission
Submitted (04-FEB-1999) Bone & Cartilage Biology, UW 2109,
SmithKline Beecham, 709 Swedeland Rd., King of Prussia, PA 19406,
USA
FEATURES
source
Location/Qualifiers
1..1739
/organism="Mus musculus"
/db_xref="taxon:10090"
/tissue_type="lung"
1..1739
/gene="Ctgfl"
242..997
/gene="Ctgfl"
/note="similar to the Mus musculus WISP-2 protein encoded
by the sequence presented in GenBank Accession Number
AF100778; putative growth factor; CTGF-L; contains IGF
binding (IGFBP), Von Willebrand Factor type C (VWC) repeat
and chromospondin type I (TSP1) domains; member of the
CCN (CTGF/Cyrl6/Nov) family; lacks the fourth
carboxy-terminal (CT) domain present in other members of
the CCN family"
/codon_start=1
/product="connective tissue growth factor-like protein
precursor"
/protein_id="AADI8058.1"
/db_xref="GI:4337060"
/translation="MRGNPLHLIAIFLCILSMVYSQCPAPACPTWPPCPGPV
LVLDGCGCRVCARLGRGCDHLVCDPSGIVCPGAGPSRGAVCLFEEDGSCVE
NGRRYLDGETFKPCRVLCRDGDTGLPCSEVRLPSWDCPPRRIQVPGRCCEV
WVCDQVMPAIPQPSAGHLSALVTPASADGPCNWSADGPCSTTCGLGIATRV
NONRCQLEIQRRLCLSRCLASRSHGWSNAF"
BASE COUNT 375 a 480 c 489 g 395 t
ORIGIN
Query Match 69.4%; Score 121.4; DB 10; Length 1739;
Best Local Similarity 84.6%; Pred. No. 3.7e-32;
Matches 148; Conservative 0; Mismatches 26; Indels 1; Gaps 1;
QY 1 AGTCAGGAACTGAGCTTGTATTTTCAGGAATGCATCTCTTAAGCACTCGCAAAAC 60
|||||
Db 1521 AGTCAGGAACTGAGCTTGTATTTGTAATAATACATCTCTTAATGCTCAAAAGC 1580
|||||
QY 61 AGGAAGGCTCCACACTCTGGAGCCAGGCGCTTCTCTTCAGCATGAGAAAGACAAG 120
|||||
Db 1581 AAG-AGGCTCCACACTTTGGCAGGCCAGGCGCTTCTCTTCAGCATGAGAGACAAG 1639
|||||
QY 121 GACACAGAGTACTCTCTCTGGAGGACTAGTCTAGCCTAGNATAAACACCCAA 175

```

```

|||||
Db 1640 AACAGTAGTAGCTCTCTCTTTGGAGGACTGGCCCGGTTTGGATAAACACCCAA 1694
|||||
RESULT 10
AC130013/c
LOCUS
DEFINITION
Rattus norvegicus clone CH230-19J17, *** SEQUENCING IN PROGRESS
***, 66 unordered pieces.
ACCESSION
AC130013
VERSION
AC130013.1 GI:22138176
KEYWORDS
HTG: HTGS_PHASE1.
SOURCE
Norway rat.
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 208219)
Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-oshan, F.R., Allen, C.,
Albrooks, S.L., Amaratunge, H.C., Are, J.R., Ayele, M., Banks, T.,
Barbaria, J., Benton, J., Bimage, K., Blankenburg, K., Bonnin, D.,
Bouck, J., Bowle, S., Brileva, M., Brown, E., Brown, M., Bryant, N.P.,
Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,
Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,
Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,
Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,
Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,
Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,
Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,
Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,
Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,
Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,
Hernandez, O., Hodgson, A., Hognes, M., Holloway, C., Hollins, B.,
Honsi, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,
Jacobson, B., Jia, F., Johnson, R., Jollivet, S., Joudah, S.,
Karlisson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,
Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulseged, H.,
Lozadó, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,
Massey, G., Mawhinney, E., McLeod, M.P., Meador, M., Mel, G., Metzker, M.,
Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, N.,
Moser, N., Nick, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
Nguyen, N., Nickerson, E., Nwokenkwo, S., Oguh, M., Okwuonu, G.,
Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
Rives, M., Rojas, A., Rojubokan, I., Rolfe, M., Ruiz, S., Savery, G.,
Scherer, S., Scott, G., Shen, H., Shooshtari, N., Sisson, I.,
Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H.,
Sutton, J., Syatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,
Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q.,
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K.,
Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G. and Gibbs, R.
Direct Submission
Unpublished
JOURNAL
TITLE
2 (bases 1 to 208219)
REFERENCE
Worley, K.C.
Direct Submission
Submitted (08-AUG-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
----- Genome Center of Medicine
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: KAYZ

```





## RESULT 11

AP003608

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

AP003608 112855 bp DNA linear HTG 21-MAR-2002  
 Oryza sativa (japonica cultivar-group) chromosome 6 clone P0007G12,  
 \*\*\* SEQUENCING IN PROGRESS \*\*\*, in ordered pieces.

AP003608

HTG; HTGS\_PHASE2.

Oryza sativa (japonica cultivar-group) (cultivar:Nipponbare) DNA,  
 clone:P0007G12.

ORGANISM

Oryza sativa (japonica cultivar-group)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

Submitted (10-MAY-2001) Takuji Sasaki, National Institute of  
 Agricultural Resources, Rice Genome Research Program; Kannondai  
 2-1-2, Tsukuba, Ibaraki 305-8602, Japan  
 (E-mail: tsasaki@affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/  
 Tel: 81-298-38-7441, Fax: 81-298-38-7468)

COMMENT

NOTE: It currently consists of 1 contigs. Gaps between the contigs  
 are represented as runs of N. The order of the pieces is believed  
 to be correct as given, however the sizes of the gaps between them  
 are based on estimates that have provided by the submitter. This  
 sequence will be replaced by the finished sequence as soon as it is  
 available and the accession number will be preserved.

\* NOTE: This is a 'working draft' sequence.

\* This sequence will be replaced

\* by the finished sequence as soon as it is available and

\* the accession number will be preserved.

FEATURES

source

1. .112855  
 Location/Qualifiers  
 /organism="Oryza sativa (japonica cultivar-group)"  
 /cultivar="Nipponbare"  
 /db\_xref="taxon:39947"  
 /chromosome="6"  
 /clone="P0007G12"

BASE COUNT 32153 a 25371 c 24380 g 30901 t 50 others

ORIGIN

Query Match 20.0%; Score 35; DB 2; Length 112855;  
 Best Local Similarity 55.3%; Pred. No. 0.78;  
 Matches 68; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 46 AAGCACTCGCAAAACAGGAAGGCTCACACCTCTGCGAGCGCCGCTTCTCTTCAGC 105

DB 81341 AGAATATCGAGTAGGGGAGGTGATCATGACCATGAACCGCAGAGCCCTTCAATGCAGT 81400

QY 106 ATGAGAAAGCAAGGACAGCAGAGTACTCTCTGAGGAGCTAGTCTAGCTAGAATA 165

DB 81401 TGGCAAGTAATTCGCCAATGCGTTGTCGCCCTCCGGAAGCATAGTACCATGGAGTA 81460

QY 166 AAC 168

DB 81461 AAC 81463

RESULT 12

AC023195

LOCUS

DEFINITION

ACCESSION

KEYWORDS

SOURCE

AC023195 143565 bp DNA linear HTG 13-JUL-2000  
 Homo sapiens clone RP11-336L21, LOW-PASS SEQUENCE SAMPLING.  
 AC023195  
 AC023195.2 GI:7417808  
 HTG; HTGS\_PHASE0.  
 Homo sapiens.

## ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 143565)

Birren,B., Linton,L., Nusbaum,C. and Lander,E.

Homo sapiens, clone RP11-336L21

Unpublished

2 (bases 1 to 143565)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,  
 Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,

Boguslavskiy,L., Boukhgalter,B., Brown,A., Burkett,G., Castle,A.,

Choepeil,Y., Collangelo,M., Collins,S., Collamore,A., Cooke,P.,

Dearellano,K., Dewar,K., Domino,M., Doyle,M., Feneator,J.,

Ferrelira,P., Fitzhugh,W., Forrest,C., Gage,D., Galagan,J.,

Gardyna,S., Grant,G., Hagos,B., Hearford,A., Horton,L.,

Howland,J.C., Johnson,R., Jones,C., Kann,L., Karas,A., Klein,J.,

Landers,T., Lehoczyk,J., Levine,R., Liu,C., Liu,G., Locke,K.,

Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,

McPheeters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J.,

Norman,C.H., O'Connor,T., O'Donnell,P., Oliver,T.M., Peterson,K.,

Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,

Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,

Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,

Tirrell,A., Vagstad,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,

Zimmer,A. and Zody,M.

Direct Submission

Submitted (09-EEB-2000) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

On Apr 5, 2000 this sequence version replaced gi:6957786.

All repeats were identified using RepeatMasker:

Smit, A.F.A. &amp; Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence\_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L5859

Center clone name: 336\_L\_21

-----

\* NOTE: This record contains 160 individual  
 \* sequencing reads that have not been assembled into  
 \* contigs. Runs of N are used to separate the reads  
 \* and the order in which they appear is completely  
 \* arbitrary. Low-pass sequence sampling is useful for  
 \* identifying clones that may be gene-rich and allows  
 \* overlap relationships among clones to be deduced.  
 \* However, it should not be assumed that this clone  
 \* will be sequenced to completion. In the event that  
 \* the record is updated, the accession number will  
 \* be preserved.

\* 1 809: contig of 809 bp in length

\* 810 909: gap of 100 bp

\* 910 1721: contig of 812 bp in length

\* 1722 1821: gap of 100 bp

\* 1822 2585: contig of 764 bp in length

\* 2586 2685: gap of 100 bp

\* 2686 3498: contig of 806 bp in length

\* 3492 3591: gap of 100 bp

\* 3592 4394: contig of 803 bp in length

\* 4395 4494: gap of 100 bp

\* 4495 5316: contig of 822 bp in length

\* 5317 5416: gap of 100 bp

\* 5417 6223: contig of 807 bp in length

\* 6224 6323: gap of 100 bp

\* 6324 7112: contig of 789 bp in length

\* 7113 7212: gap of 100 bp

\* 7213 8001: contig of 789 bp in length

\* 8002 8101: gap of 100 bp

\* 8102 8880: contig of 779 bp in length

\* 8881 8980: gap of 100 bp

\* 8981 9753: contig of 773 bp in length

\* 9754 9853: gap of 100 bp  
\* 9854 10645: contig of 792 bp in length  
\* 10646 10745: gap of 100 bp  
\* 10746 11537: contig of 792 bp in length  
\* 11538 11637: gap of 100 bp  
\* 11638 12431: contig of 794 bp in length  
\* 12432 12531: gap of 100 bp  
\* 12532 13302: contig of 771 bp in length  
\* 13303 13402: gap of 100 bp  
\* 13403 14204: contig of 802 bp in length  
\* 14205 14304: gap of 100 bp  
\* 14305 15096: contig of 792 bp in length  
\* 15097 15196: gap of 100 bp  
\* 15197 16002: contig of 806 bp in length  
\* 16003 16102: gap of 100 bp  
\* 16103 16900: contig of 798 bp in length  
\* 16901 17000: gap of 100 bp  
\* 17001 17806: contig of 806 bp in length  
\* 17807 17906: gap of 100 bp  
\* 17907 18690: contig of 784 bp in length  
\* 18691 18790: gap of 100 bp  
\* 18791 19604: contig of 814 bp in length  
\* 19605 19704: gap of 100 bp  
\* 19705 20542: contig of 838 bp in length  
\* 20543 20642: gap of 100 bp  
\* 20643 21455: contig of 813 bp in length  
\* 21456 21555: gap of 100 bp  
\* 21556 22345: contig of 790 bp in length  
\* 22346 22445: gap of 100 bp  
\* 22446 23258: contig of 813 bp in length  
\* 23259 23358: gap of 100 bp  
\* 23359 24152: contig of 794 bp in length  
\* 24153 24252: gap of 100 bp  
\* 24253 25033: contig of 781 bp in length  
\* 25034 25133: gap of 100 bp  
\* 25134 25922: contig of 789 bp in length  
\* 25923 26022: gap of 100 bp  
\* 26023 26823: contig of 801 bp in length  
\* 26824 26923: gap of 100 bp  
\* 26924 27728: contig of 805 bp in length  
\* 27729 27828: gap of 100 bp  
\* 27829 28645: contig of 817 bp in length  
\* 28646 28745: gap of 100 bp  
\* 28746 29577: contig of 832 bp in length  
\* 29578 29677: gap of 100 bp  
\* 29678 30442: contig of 765 bp in length  
\* 30443 30542: gap of 100 bp  
\* 30543 31366: contig of 824 bp in length  
\* 31367 31466: gap of 100 bp  
\* 31467 32265: contig of 799 bp in length  
\* 32266 32365: gap of 100 bp  
\* 32366 33216: contig of 851 bp in length  
\* 33217 33316: gap of 100 bp  
\* 33317 34036: contig of 720 bp in length  
\* 34037 34136: gap of 100 bp  
\* 34137 34927: contig of 791 bp in length  
\* 34928 35027: gap of 100 bp  
\* 35028 35832: contig of 805 bp in length  
\* 35833 35932: gap of 100 bp  
\* 35933 36696: contig of 764 bp in length  
\* 36697 36796: gap of 100 bp  
\* 36797 37597: contig of 801 bp in length  
\* 37598 37697: gap of 100 bp  
\* 37698 38491: contig of 794 bp in length  
\* 38492 38591: gap of 100 bp  
\* 38592 39326: contig of 735 bp in length  
\* 39327 39426: gap of 100 bp  
\* 39427 40236: contig of 810 bp in length  
\* 40237 40336: gap of 100 bp  
\* 40337 41137: contig of 801 bp in length  
\* 41138 41237: gap of 100 bp  
\* 41238 42050: contig of 813 bp in length  
\* 42051 42150: gap of 100 bp

\* 42151 42964: contig of 814 bp in length  
\* 42965 43064: gap of 100 bp  
\* 43065 43883: contig of 819 bp in length  
\* 43884 43983: gap of 100 bp  
\* 43984 44775: contig of 792 bp in length  
\* 44776 44875: gap of 100 bp  
\* 44876 45675: contig of 800 bp in length  
\* 45676 45775: gap of 100 bp  
\* 45776 46562: contig of 787 bp in length  
\* 46563 46662: gap of 100 bp  
\* 46663 47457: contig of 795 bp in length  
\* 47458 47557: gap of 100 bp  
\* 47558 48341: contig of 784 bp in length  
\* 48342 48441: gap of 100 bp  
\* 48442 49237: contig of 796 bp in length  
\* 49238 49337: gap of 100 bp  
\* 49338 50130: contig of 793 bp in length  
\* 50131 50230: gap of 100 bp  
\* 50231 51036: contig of 806 bp in length  
\* 51037 51136: gap of 100 bp  
\* 51137 51929: contig of 793 bp in length  
\* 51930 52029: gap of 100 bp  
\* 52030 52844: contig of 815 bp in length  
\* 52845 52944: gap of 100 bp  
\* 52945 53752: contig of 808 bp in length  
\* 53753 53852: gap of 100 bp  
\* 53853 54670: contig of 818 bp in length  
\* 54671 54770: gap of 100 bp  
\* 54771 55560: contig of 790 bp in length  
\* 55561 55660: gap of 100 bp  
\* 55661 56464: contig of 804 bp in length  
\* 56465 56564: gap of 100 bp  
\* 56565 57370: contig of 806 bp in length  
\* 57371 57470: gap of 100 bp  
\* 57471 58244: contig of 774 bp in length  
\* 58245 58344: gap of 100 bp  
\* 58345 59162: contig of 818 bp in length  
\* 59163 59262: gap of 100 bp  
\* 59263 60007: contig of 745 bp in length  
\* 60008 60107: gap of 100 bp  
\* 60108 60942: contig of 835 bp in length  
\* 60943 61042: gap of 100 bp  
\* 61043 61837: contig of 795 bp in length  
\* 61838 61937: gap of 100 bp  
\* 61938 62743: contig of 806 bp in length  
\* 62744 62843: gap of 100 bp  
\* 62844 63636: contig of 793 bp in length  
\* 63637 63736: gap of 100 bp

Query Match 19.9% Score 34.8; DB 2; Length 143565;  
Best Local Similarity 65.4%; Pred. No.: 0.96;  
Matches 51; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 43 CTTAAGCAGCTCGGAAACAGGAGGCTCCACCTCTGGAGCGGCGCTTTCTCTTC 102  
DB 114635 CTGCGTAGGCAGAAACTGGAATCGTGCGGAGGCGCTTCCTTTTG 114694  
QY 103 AGCATGAGAAAGCAAGG 120  
DB 114695 GTCAGGGGAAAGTCCAGG 114712

RESULT 13  
AL671848/c  
LOCUS AL671848 bp DNA linear ROD 09-JUL-2002  
DEFINITION Mouse DNA sequence from clone RP23-110K5 on chromosome X, complete  
sequence.  
ACCESSION AL671848  
VERSION AL671848.7 GI:21727394  
KEYWORDS HTG.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

COMMENT

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 190564)  
Tracey, A.  
Direct Submission  
Submitted (09-JUL-2002) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
humquery@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk  
On Jul 10, 2002 this sequence version replaced gi:21655363.  
----- Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: <http://www.sanger.ac.uk>  
Contact: [humquery@sanger.ac.uk](mailto:humquery@sanger.ac.uk)  
-----  
During sequence assembly data is compared from overlapping clones.  
Where differences are found these are annotated as variations  
together with a note of the overlapping clone name. Note that the  
variation annotation may not be found in the sequence submission  
corresponding to the overlapping clone, as we submit sequences with  
only a small overlap as described above.  
This sequence was finished as follows unless otherwise noted: all  
regions were either double-stranded or sequenced with an alternate  
chemistry or covered by high quality data (i.e., phred quality >=  
30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by at least  
one plasmid subclone or more than one M13 subclone; and the  
assembly was confirmed by restriction digest. The following  
abbreviations are used to associate primary accession numbers given  
in the feature table with their source databases: Em; EMBL; SW;  
SWISSPROT; Tr; TREMBL; Wp; WORMPEP; Information on the WORMPEP  
database can be found at  
[http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) RP23-110K5 is  
from the R6C1-23 Mouse PAC Library  
constructed by the group of Pieter de Jong.  
For further details see <http://www.chori.org/bacpac/home.htm>  
VECTOR: pBACE3.6.

| FEATURES                 | source   |
|--------------------------|--|
| Location/Qualifiers      | i..190964  |
| /organism="Mus musculus" |  |
| /db_xref="taxon:10090"   |  |
| /chromosome="X"          |  |
| /clone="RP23-110K5"      |  |
| /clone_lib="RPCI-23"     |  |
| BASE COUNT               | 61228 a 37923 c 37578 g 54235 t  |
| ORIGIN                   |  |
| Query Match              | 19.9%; Score 34.8; DB 10;  |
| Best Local Similarity    | 58.8%; Pred. No. 1;  |
| Matches                  | 60; Conservative 0; Mismatches 42; Indels 0; Gaps 0;                   |
| QY                       | 36 CACATCTCTTAAGCACTGCCAAACAGGAAGGCTCCACACCCTTCGGCAGGCAGGCGCTT 95<br>  |
| Dd                       | 54490 CACAATCTCTAACCATCTCCCACCTGGGCATCTCTCCAAGTCTCTCTAGGTAGTGCTT 54431 |
| QY                       | 96 TCTTTTCAGCATGAGAAGACGAAGGCACAGCAGACTACTCTC 137<br>                  |
| Dd                       | 54430 CCTCTGCCACTGAAGACACAGCAAGCACACCCTGGAACCATCTC 54389<br>           |

|            |  |
|------------|--|
| RESULT 14  |  |
| AC094291   |  |
| LOCUS      | 233269 bp DNA linear HTG 09-JUL-2002   |
| DEFINITION | Rattus norvegicus clone CH230-3H24, *** SEQUENCING IN PROGRESS ***,<br>60 unordered pieces.  |
| ACCESSION  | AC094291   |
| VERSION    | AC094291.3 GI:21716642   |
| KEYWORDS   | HTG; HTGS_PHASE1.  |
| SOURCE     | Norway rat.  |
| ORGANISM   | Rattus norvegicus  |
|            | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;<br>Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; |
| REFERENCE  | 1 (bases 1 to 233269)  |

## AUTHORS

Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C., Alsbark, S.L., Anaratunga, H.C., Are, J.R., Ayele, M., Banks, T., Bartholais, J., Benton, J., Bimge, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowle, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P., Buahy, C., Burch, P., Burrett, C., Burrell, K.L., Byrd, N.C., Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorielli, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hoques, M., Holloway, C., Hollins, B., Homi, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, O., King, L., Korvan, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Lounsegh, H., Lozaro, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhiney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Monabbat, N., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, M., Nguyen, N., Nguyen, N., Nickerson, E., Nwokwkw, S., Oguh, M., Okwuonu, G., Ogunyeye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Qulles, M., Ren, Y., Rivers, M., Rojas, A., Rojbum, I., Rolfe, M., Ruiz, S., Savery, G., Scherek, S., Scott, G., Shen, H., Shoohtari, N., Sison, I., Sodergren, E., Sonaik, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstein, G., and Gibbs, R.

Direct Submission  
2 (bases 1 to 233269)  
Unpublished  
Worley, K.C.

Submitted (17-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 233269)  
Worley, K.C.

Direct Submission  
Submitted (09-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On Jul 9, 2002 this sequence version replaced qi:17941008.

```

Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information -----
Center project name: GAIL
Center clone name: CH230-3H24
----- Summary Statistics -----
Sequencing vector: Plasmid;
Sequencing vector: M13;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 175720 bases at least Q40
Consensus quality: 180089 bases at least Q30
Consensus quality: 183792 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length

```

\* NOTE: Estimated insert size may differ from sequence length

(see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).

\* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 60 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

\* 1 1068: contig of 1068 bp in length  
 \* 1069 1168: gap of unknown length  
 \* 1169 2363: contig of 1195 bp in length  
 \* 2363 2463: gap of unknown length  
 \* 2463 2682: contig of 1219 bp in length  
 \* 2682 3683: gap of unknown length  
 \* 3683 3783: contig of 1115 bp in length  
 \* 3783 4898: gap of unknown length  
 \* 4898 6250: contig of 1253 bp in length  
 \* 6250 6351: gap of unknown length  
 \* 6351 7467: contig of 1117 bp in length  
 \* 7467 7568: gap of unknown length  
 \* 7568 8743: contig of 1175 bp in length  
 \* 8743 8843: gap of unknown length  
 \* 8843 10039: contig of 1197 bp in length  
 \* 10039 10139: gap of unknown length  
 \* 10139 11793: contig of 1654 bp in length  
 \* 11793 11894: gap of unknown length  
 \* 11894 13371: contig of 1478 bp in length  
 \* 13371 13471: gap of unknown length  
 \* 13471 14864: contig of 1393 bp in length  
 \* 14864 14964: gap of unknown length  
 \* 14964 14965: contig of 1496 bp in length  
 \* 14965 16461: gap of unknown length  
 \* 16461 16560: contig of 1103 bp in length  
 \* 16560 17664: gap of unknown length  
 \* 17664 19281: contig of 1518 bp in length  
 \* 19281 19382: gap of unknown length  
 \* 19382 21112: contig of 1731 bp in length  
 \* 21112 21213: gap of unknown length  
 \* 21213 22833: contig of 1621 bp in length  
 \* 22833 24214: contig of 1281 bp in length  
 \* 24214 24314: gap of unknown length  
 \* 24314 25527: contig of 1213 bp in length  
 \* 25527 25628: gap of unknown length  
 \* 25628 27220: contig of 1593 bp in length  
 \* 27220 27320: gap of unknown length  
 \* 27320 28926: contig of 1606 bp in length  
 \* 28926 29026: gap of unknown length  
 \* 29026 30363: contig of 1337 bp in length  
 \* 30363 30464: gap of unknown length  
 \* 30464 31665: contig of 1202 bp in length  
 \* 31665 33349: gap of unknown length  
 \* 33349 33449: contig of 1584 bp in length  
 \* 33449 36317: contig of 2868 bp in length  
 \* 36317 36418: gap of unknown length  
 \* 36418 37778: contig of 1361 bp in length  
 \* 37778 39831: contig of 1973 bp in length  
 \* 39831 39951: gap of unknown length  
 \* 39951 42802: contig of 2851 bp in length  
 \* 42802 42902: gap of unknown length  
 \* 42902 45682: contig of 2780 bp in length  
 \* 45682 45783: gap of unknown length  
 \* 45783 49268: contig of 3486 bp in length  
 \* 49268 51428: gap of unknown length  
 \* 51428 51529: contig of 2060 bp in length  
 \* 51529 54571: gap of unknown length  
 \* 54571 54671: contig of 3043 bp in length  
 \* 54671 54672: gap of unknown length  
 \* 54672 57546: contig of 2875 bp in length  
 \* 57546 57646: gap of unknown length

\* 57647 60741: contig of 3094 bp in length  
 \* 60741 63916: gap of unknown length  
 \* 63916 64016: contig of 3076 bp in length  
 \* 64016 67001: gap of unknown length  
 \* 67001 70219: contig of 2984 bp in length  
 \* 70219 70319: gap of unknown length  
 \* 70319 72950: contig of 3119 bp in length  
 \* 72950 73051: gap of unknown length  
 \* 73051 76802: contig of 2631 bp in length  
 \* 76802 76902: gap of unknown length  
 \* 76902 79990: contig of 3752 bp in length  
 \* 79990 80090: gap of unknown length  
 \* 80090 82745: contig of 3088 bp in length  
 \* 82745 82845: gap of unknown length  
 \* 82845 87725: contig of 2655 bp in length  
 \* 87725 87825: gap of unknown length  
 \* 87825 91178: contig of 4880 bp in length  
 \* 91178 91278: gap of unknown length  
 \* 91278 94534: contig of 3353 bp in length  
 \* 94534 94634: contig of 3255 bp in length  
 \* 94634 99871: gap of unknown length  
 \* 99871 99971: contig of 5238 bp in length  
 \* 99971 103743: gap of unknown length  
 \* 103743 103842: contig of 3771 bp in length  
 \* 103842 108877: gap of unknown length  
 \* 108877 108977: contig of 5035 bp in length  
 \* 108977 114019: gap of unknown length  
 \* 114019 114119: contig of 5042 bp in length  
 \* 114119 119231: gap of unknown length  
 \* 119231 119332: contig of 5112 bp in length  
 \* 119332 125649: gap of unknown length  
 \* 125649 125749: contig of 6318 bp in length  
 \* 125749 132687: gap of unknown length  
 \* 132687 132787: contig of 6938 bp in length  
 \* 132787 139741: gap of unknown length  
 \* 139741 139841: contig of 6954 bp in length  
 \* 139841 146349: gap of unknown length  
 \* 146349 146350: contig of 6508 bp in length  
 \* 146350 155325: gap of unknown length  
 \* 155325 155425: contig of 8876 bp in length  
 \* 155425 155326: gap of unknown length

Query Match 19.9%; Score 34.8; DB 2; Length 233269;  
 Best Local Similarity 62.8%; Pred. No. 1.1;  
 Matches 54; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 39 ATCTCTTAAGCACTCGCAAAACAGGAGGCTCCACACCTCTGCGAGCCGCGCTTTCT 98  
 Db 95327 ATCTCCAAGGTAACGTTATTAAAGAAAGCTCCAAACAGCTGTGCCGTCATCGTCATTCT 95386  
 QY 99 CTTTCAGCATGAGAAAGCAGGAGCA 124  
 Db 95387 TCTCAACAGATAAAGACAGGTGCCA 95412

# RESULT 15

AC068103/c  
 LOCUS AC068103 144233 bp DNA linear HTG 14-JUN-2000  
 DEFINITION Homo sapiens chromosome 4 clone RP11-18114 map 4, WORKING DRAFT  
 SEQUENCE, 38 unordered pieces.  
 AC068103  
 VERSION AC068103.2 GI:8516095  
 KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 144233)  
 Birtten, B., Linton, L., Nusbaum, C. and Lander, E.  
 TITLE Homo sapiens chromosome 4, clone RP11-18114  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 144233)

## AUTHORS

Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N., Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F., Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G., Campolano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L., Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R., Meltrin,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and Zody,M.

TITLE  
JOURNAL

Direct Submission

Submitted (28-APR-2000) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

On Jun 14, 2000 this sequence version replaced gi:7658404.

All repeats were identified using RepeatMasker:

Smit, A.F.A. &amp; Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)

----- Project Information

Center project name: L7833

Center clone name: L81\_I\_4

----- Summary Statistics

Sequencing vector: M13; M7815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 125189 bases at least Q40

Consensus quality: 133535 bases at least Q30

Consensus quality: 137574 bases at least Q20

Insert size: 160000; agarose-1p

Quality coverage: 3.1 in Q20 bases; agarose-1p

Quality coverage: 3.5 in Q20 bases; sum-of-contigs

-----

\* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 38 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

i 1144: contig of 1144 bp in length  
 \* 1145 1244: gap of 100 bp  
 \* 1245 2299: contig of 1055 bp in length  
 \* 2300 2399: gap of 100 bp  
 \* 2400 3022: contig of 1423 bp in length  
 \* 3023 3922: gap of 100 bp  
 \* 3923 5190: contig of 1268 bp in length  
 \* 5191 5290: gap of 100 bp  
 \* 5291 5723: contig of 433 bp in length  
 \* 5724 5823: gap of 100 bp  
 \* 5824 7679: contig of 1856 bp in length  
 \* 7680 7779: gap of 100 bp  
 \* 7780 9186: contig of 1407 bp in length  
 \* 9187 9286: gap of 100 bp  
 \* 9287 11735: contig of 2449 bp in length  
 \* 11736 11835: gap of 100 bp  
 \* 11836 13249: contig of 1414 bp in length

FEATURES  
Source

1. 144233  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="4"  
 /map="4"  
 /clone="RP11-181I4"  
 /clone.lib="RPC1-11 Human Male BAC"  
 1..1144  
 /note="assembly\_fragment"  
 1245..2299  
 /note="assembly\_fragment"  
 2400..3822  
 /note="assembly\_fragment"  
 3923..5190  
 /note="assembly\_fragment"

\* 13250 13349: gap of 100 bp  
 \* 13350 14701: contig of 1352 bp in length  
 \* 14702 14801: gap of 100 bp  
 \* 14802 16706: contig of 1905 bp in length  
 \* 16707 16806: gap of 100 bp  
 \* 16807 18769: contig of 1963 bp in length  
 \* 18770 18869: gap of 100 bp  
 \* 18870 20539: contig of 1670 bp in length  
 \* 20540 20639: gap of 100 bp  
 \* 20640 23269: contig of 2630 bp in length  
 \* 23270 23369: gap of 100 bp  
 \* 23370 25610: contig of 2241 bp in length  
 \* 25611 25710: gap of 100 bp  
 \* 25711 27748: contig of 2038 bp in length  
 \* 27749 27848: gap of 100 bp  
 \* 27849 29177: contig of 1329 bp in length  
 \* 29178 29277: gap of 100 bp  
 \* 29278 30870: contig of 1593 bp in length  
 \* 30871 30970: gap of 100 bp  
 \* 30971 34125: contig of 3155 bp in length  
 \* 34126 34225: gap of 100 bp  
 \* 34226 36923: contig of 2698 bp in length  
 \* 36924 37023: gap of 100 bp  
 \* 37024 39589: contig of 2566 bp in length  
 \* 39590 39689: gap of 100 bp  
 \* 39690 42352: contig of 2663 bp in length  
 \* 42353 42452: gap of 100 bp  
 \* 42453 45755: contig of 3303 bp in length  
 \* 45756 45855: gap of 100 bp  
 \* 45856 49062: contig of 3207 bp in length  
 \* 49063 49162: gap of 100 bp  
 \* 49163 51707: contig of 2545 bp in length  
 \* 51708 51807: gap of 100 bp  
 \* 51808 56155: contig of 4348 bp in length  
 \* 56156 56255: gap of 100 bp  
 \* 56256 61528: contig of 5273 bp in length  
 \* 61529 61628: gap of 100 bp  
 \* 61629 66907: contig of 5279 bp in length  
 \* 66908 67007: gap of 100 bp  
 \* 67008 72627: contig of 5620 bp in length  
 \* 72628 72727: gap of 100 bp  
 \* 72728 77672: contig of 4945 bp in length  
 \* 77673 77772: gap of 100 bp  
 \* 77773 85211: contig of 7439 bp in length  
 \* 85212 85311: gap of 100 bp  
 \* 85312 92800: contig of 7489 bp in length  
 \* 92801 92900: gap of 100 bp  
 \* 92901 99694: contig of 6794 bp in length  
 \* 99695 99794: gap of 100 bp  
 \* 99795 108698: contig of 8904 bp in length  
 \* 108699 108798: gap of 100 bp  
 \* 108799 116661: contig of 7863 bp in length  
 \* 116662 116761: gap of 100 bp  
 \* 116762 128568: contig of 11807 bp in length  
 \* 128569 128668: gap of 100 bp  
 \* 128669 136426: contig of 7758 bp in length  
 \* 136427 136526: gap of 100 bp  
 \* 136527 144233: contig of 7707 bp in length.

misc\_feature  
 misc\_feature  
 misc\_feature  
 misc\_feature

misc\_feature /note="assembly\_fragment"  
5291..5723  
/note="assembly\_fragment"  
clone\_end:T7  
vector\_side:right"  
5824..7679  
/note="assembly\_fragment"  
7780..9186  
/note="assembly\_fragment"  
9287..11735  
/note="assembly\_fragment"  
11836..13249  
/note="assembly\_fragment"  
13350..14701  
/note="assembly\_fragment"  
14802..16706  
/note="assembly\_fragment"  
16807..18769  
/note="assembly\_fragment"  
18870..20539  
/note="assembly\_fragment"  
20640..23269  
/note="assembly\_fragment"  
23370..25610  
/note="assembly\_fragment"  
25711..27748  
/note="assembly\_fragment"  
27849..29177  
/note="assembly\_fragment"  
29278..30870  
/note="assembly\_fragment"  
30971..34125  
/note="assembly\_fragment"  
34226..36923  
/note="assembly\_fragment"  
37024..39589  
/note="assembly\_fragment"  
39690..42352  
/note="assembly\_fragment"

Query Match 19.78; Score 34.4; DB 2; Length 144233;  
Best Local Similarity 54.88; Pred. No. 1.3;  
Matches 68; Conservative 0; Mismatches 56; Indels 0; Gaps 0;  
Qy 9 AACTTGAGCTTGTATTTTCAGGAATGCACATCTCTTAAGCACTCGCAAAACAGGAAGGC 68  
Db 72574 ACCTGAAGCTCTGTGTTTCAGGGATCAGACTTCCCTCACAGTTACCAGAAAGGCCAGGT 72515  
Qy 69 TCCACACCTCTGGCAGGCCAGGCCCTTCTCTTCAGCATGAGAAAGACAGGACAGCAG 128  
Db 72514 AACACAACCTCAATAAACTCAGGGACTTAGTGCTCCTCTGTGGAAACAAATTGACCAC 72455  
Qy 129 AGTA 132  
Db 72454 AGAA 72451

Search completed: July 29, 2003, 00:48:37  
Job time : 491.369 secs





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 23, 2003, 09:03:08 ; Search time 46.3409 Seconds  
(without alignments)  
8504.365 Million cell updates/sec

Title: US-10-010-408-1\_COPY\_1534\_1708

Perfect score: 175

Sequence: 1 AGTCCAGGAAGCTTACGCTTT.....GCCTAGAGTAACACCCAAA 175

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_101002.\*

1: /SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 1          | 175   | 100.0       | 1708   | 20 | AAZ07516    |
| 2          | 126.2 | 72.1        | 1734   | 20 | AAZ76488    |
| 3          | 125.8 | 71.9        | 439    | 24 | ABL59575    |
| 4          | 57    | 32.6        | 65     | 24 | ABN30189    |
| 5          | 30.8  | 17.6        | 463    | 22 | AAK54634    |
| 6          | 30.4  | 17.4        | 2155   | 23 | ABL11587    |
| 7          | 30.4  | 17.4        | 14091  | 23 | ABL11586    |
| 8          | 30    | 17.1        | 330    | 22 | AAI91961    |
| 9          | 30    | 17.1        | 422    | 22 | AAF64359    |

|    |      |      |         |    |           |                    |
|----|------|------|---------|----|-----------|--------------------|
| 10 | 30   | 17.1 | 483     | 22 | ABA42323  | Human breast cell  |
| 11 | 30   | 17.1 | 483     | 22 | ABA52750  | Human foetal liver |
| 12 | 30   | 17.1 | 483     | 22 | ABA22534  | Probe #1000 for ge |
| 13 | 30   | 17.1 | 483     | 22 | AAK01005  | Human brain exres  |
| 14 | 30   | 17.1 | 483     | 22 | AAK26462  | Human bone marrow  |
| 15 | 30   | 17.1 | 483     | 22 | AAI11091  | Probe #1024 for ge |
| 16 | 30   | 17.1 | 483     | 22 | AAI32356  | Probe #1042 used t |
| 17 | 30   | 17.1 | 483     | 22 | AAI01010  | Probe #1001 used t |
| 18 | 30   | 17.1 | 483     | 22 | ABSO1052  | Human genome-deriv |
| 19 | 30   | 17.1 | 1096    | 22 | ABA47464  | Human breast cell  |
| 20 | 30   | 17.1 | 1096    | 22 | ABA65355  | Human foetal liver |
| 21 | 30   | 17.1 | 1096    | 22 | ABA32452  | Probe #10918 for g |
| 22 | 30   | 17.1 | 1096    | 22 | AAK13765  | Human brain exres  |
| 23 | 30   | 17.1 | 1096    | 22 | AAK39508  | Human bone marrow  |
| 24 | 30   | 17.1 | 1096    | 22 | AAI20317  | Probe #10250 for g |
| 25 | 30   | 17.1 | 1096    | 22 | AAI45522  | Probe #14208 used  |
| 26 | 30   | 17.1 | 1096    | 22 | AAI06020  | Probe #6011 used t |
| 27 | 30   | 17.1 | 1096    | 22 | ABSI3598  | Human genome-deriv |
| 28 | 30   | 17.1 | 1662    | 21 | AAF22351  | Human secreted pro |
| 29 | 30   | 17.1 | 2016    | 22 | ABA08988  | Human glucocortico |
| 30 | 29.8 | 17.0 | 2292    | 21 | AACT4236  | Human secreted pro |
| 31 | 29.8 | 17.0 | 2304    | 21 | AAC74249  | Human secreted pro |
| 32 | 29.8 | 17.0 | 1230025 | 20 | AAK91990  | Nucleotide sequenc |
| 33 | 29.4 | 16.8 | 6236    | 24 | AAI31108  | Human transcriptio |
| 34 | 29.4 | 16.8 | 25000   | 24 | ABL68348  | Kidney cancer rela |
| 35 | 29.2 | 16.7 | 1721    | 22 | AAF68139  | Murine cDNA encodi |
| 36 | 29.2 | 16.7 | 1984    | 19 | AAV01596  | Tobacco laccase cl |
| 37 | 29   | 16.6 | 1568    | 24 | ABK28657  | Human cDNA encodin |
| 38 | 29   | 16.6 | 1718    | 21 | AACT76916 | Human ORFX ORF2471 |
| 39 | 29   | 16.6 | 2658    | 23 | AAST1055  | DNA encoding novel |
| 40 | 28.8 | 16.5 | 274     | 24 | ABN76962  | Human ORF1909 cDNA |
| 41 | 28.8 | 16.5 | 5726    | 23 | ABL02959  | Drosophila melanog |
| 42 | 28.8 | 16.5 | 8343    | 23 | ABL02958  | Drosophila melanog |
| 43 | 28.6 | 16.3 | 3054    | 22 | AAH98316  | Human ESP-derived  |
| 44 | 28.6 | 16.3 | 4758    | 23 | AAST2577  | DNA encoding novel |
| 45 | 28.6 | 16.3 | 5955    | 23 | AA565525  | DNA encoding novel |

#### ALIGNMENTS

##### RESULT 1

AAZ07516  
ID AAZ07516 standard; cDNA: 1708 BP.

XX AC AAZ07516;

XX DT 26-NOV-1999 (first entry)

XX DE Rat HICP polypeptide encoding cDNA.

XX DE Heparin-induced CCN-like protein; HICP; cell-associated activity; ss;  
KW Cardiovascular disorder; aberrant cell proliferation; fibrotic disorder.

XX OS Rattus sp.

XX PN WO9947556-A2.

XX PD 23-SEP-1999.

XX PF 18-MAR-1999; 99WO-US05999.

XX PR 19-MAR-1998; 98US-0044273.

XX PA (TUFT ) TUFTS COLLEGE.

XX PI Castellot JJ;

XX DR WPI; 1999-562060/47.

XX DR P-PSDB; AAY27434.

XX PT Nucleic acid sequences encoding rat heparin-induced CCN-like protein, used in methods to identify modulators or in diagnostic applications

XX PS Claim 2; Fig 1: 108pp: English.

XX CC This cDNA encodes a rat heparin-induced CCN-like protein (HICP) protein.

CC Agents that stimulate or inhibit HICP protein activity or expression,

CC antisense HICP nucleic acid molecules and HICP antibodies, can be used to

CC modulate cell-associated activity. HICP modulators can be used to treat

CC disorders characterized by aberrant HICP protein activity or expression.

CC Probes capable of hybridizing to HICP mRNA or antibodies specific for

CC HICP can be used to detect HICP activity in a biological sample. HICP

CC can be used to treat disorders, such as a cardiovascular or fibrotic

CC disorder, characterized by aberrant cell proliferation.

XX SQ Sequence 1708 BP; 362 A; 486 C; 478 G; 382 T; 0 other;

Query Match 100.0%; Score 175; DB 20; Length 1708;

Best Local Similarity 100.0%; Pred. No. 1.4e-50;

Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGTCCAGGAACTTGAGCTTTGTATTTCAGGAATGCACATCTCTTAAGCACTCGCAAAAC 60

DB 1534 AGTCCAGGAACTTGAGCTTTGTATTTCAGGAATGCACATCTCTTAAGCACTCGCAAAAC 1593

QY 61 AGGAAGGCTCCACACCTCTGGCAGGCGGCGCTTCTTCAGCATGAGAAGACAAGG 120

DB 1594 AGGAAGGCTCCACACCTCTGGCAGGCGGCGCTTCTTCAGCATGAGAAGACAAGG 1653

QY 121 GACAGCAGAGTACTCTCTCTGGAGCACTAGTCTAGCCCTAGATAAACACCCAAA 175

DB 1654 GACAGCAGAGTACTCTCTCTGGAGCACTAGTCTAGCCCTAGATAAACACCCAAA 1708

RESULT 2

AAAX76488

ID AAX76488 standard; DNA; 1734 BP.

XX AC AAX76488;

XX DT 06-AUG-1999 (first entry)

XX DE Mouse WISP-2 protein nucleotide sequence SEQ ID NO:17.

XX WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;

KW connective tissue growth factor; cancer; melanoma; arteriosclerosis;

KW leukaemia; lymphoid malignancy; haematopoiesis-related disorder;

KW tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;

KW kidney disorder; bone-related disorder; osteoporosis; trauma; burn;

KW connective tissue disorder; catabolic state; inflammation;

KW testicular-related disorder; angiogenesis; immunological disorder; ss.

XX OS Mus sp.

XX WO9921998-Al.

XX PD 06-MAY-1999.

XX PF 29-OCT-1998; 98WO-US22991.

XX PR 14-APR-1998; 98US-0081695.

XX PR 29-OCT-1997; 97US-0063704.

XX PR 03-FEB-1998; 98US-0073612.

XX (GETH ) GENENTECH INC.

XX Botstein DA, Cohen RL, Goddard A, Gurney AL, Hillan K;

PI Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WI;

XX WPI; 1999-337420/28.

DR P-PSDB; AAY17651.

XX New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3

XX Example 2; Page 178-179; 284pp: English.

XX CC The present invention describes Wnt-1 induced secreted polypeptides,

CC WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2

CC and WISP-3 have homology to connective tissue growth factor (CTGF).

CC Products from the present invention can be used to treat WISP-related

CC disorders such as breast, ovarian, and colon cancer or melanoma. The

CC products can be used to treat arteriosclerosis. The products can also be

CC used to treat other diseases e.g. benign and malignant tumours,

CC leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,

CC hypothalamic and other glandular, macrophagal, epithelial, stromal, and

CC blastocoele disorders, haematopoiesis-related disorders, tissue-growth

CC disorders, skin disorders, desmoplasia, fibrotic lesions, kidney

CC disorders, bone-related disorders such as osteoporosis, trauma such as

CC burns, incisions, and other wounds, connective tissue disorders,

CC catabolic states, testicular-related disorders, and inflammatory,

CC angiogenic and immunologic disorders including arteriosclerosis. The

CC products can also be used for detection and diagnosis especially of

CC individuals with neoplastic cell growth or proliferation. The products

CC can be used in the production of transgenic or knock-out animals.

CC Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing

CC cells.

XX SQ Sequence 1734 BP; 355 A; 491 C; 495 G; 393 T; 0 other;

Query Match 72.1%; Score 126.2; DB 20; Length 1734;

Best Local Similarity 86.3%; Pred. No. 1.4e-33;

Matches 151; Conservative 0; Mismatches 23; Indels 1; Gaps 1;

QY 1 AGTCCAGGAACTTGAGCTTTGTATTTCAGGAATGCACATCTCTTAAGCACTCGCAAAAC 60

DB 1536 AGTCCAGGAACTTGAGCTTTGTATTTCAGGAATGCACATCTCTTAAGCACTCGCAAAAC 1595

QY 61 AGGAAGGCTCCACACCTCTGGCAGGCGGCGCTTCTTCAGCATGAGAAGACAAGG 120

DB 1596 AG-AGGCTCCACACTCTGGCAGGCGGCGCTTCTTCAGCATGAGAAGACAAGG 1654

QY 121 GACAGCAGAGTACTCTCTCTGGAGCACTAGTCTAGCCCTAGATAAACACCCAAA 175

DB 1655 AACAGTAGAGTAGCCCTCTCTGGAGCACTGGCCGCTCTGGAATAAACACCCAAA 1709

RESULT 3

ABL59575

ID ABL59575 standard; DNA; 439 BP.

XX AC ABL59575;

XX DT 17-JUL-2002 (first entry)

XX DE Rat OST23 gene fragment SEQ ID NO:23.

XX Rat; OST; osteoregenerative; parathyroid hormone; tibiae; osteopathic;

KW gene therapy; parathyroid hormone receptor ligand; bone disorder;

KW bone formation disorder; bone resorption disorder; osteoporosis;

KW osteopenia; osteopetrosis; gene; ds.

XX OS Rattus sp.

XX WO200224943-A2.

XX PD 28-MAR-2002.

XX PF 19-SEP-2001; 2001WO-US29548.

XX PR 19-SEP-2000; 2000US-233579P.

XX (CURA-) CURAGEN CORP.

PA (GLAX ) GLAXO GROUP LTD.

PI Horesovsky GJ, Noel US, Raha D;

XX WPI; 2002-401989/43.

DR

PT Identifying parathyroid hormone receptor ligands and osteoregenerative  
PT agents involves detecting the expression of nucleic acids which are  
PT regulated by parathyroid hormone  
PS  
PS Claim 45; Page 22; 90pp; English.  
XX  
XX The present invention describes a method (M1) for identifying parathyroid  
CC hormone receptor ligands (I) and osteoregenerative agents by contacting a  
CC test cell population (CP) comprising cells expressing nucleic acid  
CC sequences (S) of OST1-47 and 48, with a test agent, measuring nucleic  
CC acid sequence expression, comparing it with reference CP and identifying  
CC the ligand and agent by the difference in expression levels. Also  
CC described is a method (M2) for treating a bone disorder in a subject by  
CC administering to the subject an agent that modulates the expression or  
CC activity of (S). (I) have osteopathic activities, and can be used to  
CC modulate the expression of OST 1-48, and can also be used in gene  
CC therapy. (M1) is useful for identifying parathyroid hormone receptor  
CC ligands (I) or osteoregenerative agents. OST gene sequence can be used  
CC for assessing the osteoregenerative activity of a test agent in a  
CC subject, and for diagnosing or determining the susceptibility to bone  
CC disorder, and assessing efficacy of a treatment of a bone disorder in a  
CC subject, human or rodent. (M2) is useful for treating a bone disorder  
CC including bone formation disorder or bone resorption disorder (e.g.  
CC osteoporosis, osteopenia and osteopetrosis). OST polypeptides are useful  
CC as immunogens to raise anti-OST antibodies (II). Parathyroid hormone  
CC receptor ligands (I) identified by (M1) are useful for treating bone  
CC disorders. The present sequence represents an OST23 gene fragment  
CC isolated from rat tibiae, from the present invention.  
XX  
XX Sequence 439 BP; 128 A; 108 C; 99 G; 104 T; 0 other;

Query Match 71.9%; Score 125.8; DB 24; Length 439;  
Best Local Similarity 96.8%; Pred. No. 1.1e-33;  
Matches 150; Conservative 0; Mismatches 2; Indels 3; Gaps 2;  
Qy 1 AGTCAGGAAGTTCAGCTTTGTTATTTTCAGGAATGCACATCTCTTAAGCACTCGCAAAAC 60  
Db 285 AGTCAGGAAGTTCAGCTTTGTTATTTTCAGGAATGCACATCTCTTAAGCACTCGCAAAAC 344  
Qy 61 AGGAAGGTCACACCTCTGTGCGAGGCCGCGCTTTCTTTCAGCATGAGAAACAGG 120  
Db 345 AGGAAGGTCACACCTCTGTGCGAGGCCGCGCTTTCTTTCAGCATGAGAAACAGG 404  
Qy 121 G--ACACAGAGTAC-TCTCTCTGGAGGACTAGT 152  
Db 405 GGACCAGCAGAGTACTTTTCTCTGAGGACTAGT 439

RESULT 4  
ABN30189  
ID ABN30189 standard; DNA; 65 BP.  
XX  
XX ABN30189;  
AC  
AC  
DT 15-JUL-2002 (first entry)  
XX  
XX Rat spliced transcript detection oligonucleotide SEQ ID NO:2937.  
DE Human; mouse; rat; splice transcript; detection; RNA transcript;  
KW splice variant; transcriptome; oligonucleotide library; ss.  
XX  
XX Rattus norvegicus.  
OS  
OS WO200210449-A2.  
FN  
FN 07-FEB-2002.  
XX  
XX 20-JUL-2001; 2001WO-IB01903.  
PF  
XX 28-JUL-2000; 2000US-221607P.  
PR  
PR 02-MAY-2001; 2001US-287724P.  
XX  
XX (COMP-) COMPUGEN INC.

XX Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;  
PI WPI; 2002-257383/30.  
XX  
XX New oligonucleotide libraries comprising oligonucleotides which  
PT selectively hybridize to mRNAs transcribed from a transcription unit of  
PT a genome, useful for detecting tissue-, pathology-, and  
PT developmental-specific genes  
XX  
XX Example 1; SEQ ID 2937; 47pp; English.  
XX  
XX The present invention describes oligonucleotide libraries for detecting  
CC messenger RNAs that populate a (sub-)transcriptome, where the  
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple  
CC transcription units that populate a genome. The library comprises  
CC several oligonucleotides, each capable of hybridising selectively to a  
CC set of messenger RNAs transcribed from a given transcription unit of  
CC the genome, which encodes one or more messenger RNA splice variants.  
CC The oligonucleotide libraries are useful for detecting mRNAs from a  
CC biological sample, in expression profiling studies, in qualitatively or  
CC quantitatively characterising the corresponding transcriptome, and in  
CC detecting RNA transcripts and splice variants of human or animal  
CC transcriptomes. The libraries may also be used as specialised mini  
CC libraries to detect transcripts of a sub-transcriptome under a  
CC particular biological or pathological state, and so allowing the  
CC detection of tissue- and pathology-specific genes such as those genes  
CC only expressed in specific tissue under a specific pathological  
CC condition; to detect developmental specific genes; and to detect RNA  
CC transcripts and splice variants of a transcriptome of a patient suffering  
CC from a particular disorder. ABN27253 to ABN59389 represent  
CC oligonucleotide sequences from rats, humans and mice, which are used in  
CC the exemplification of the present invention.  
CC N.B. The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX Sequence 65 BP; 21 A; 20 C; 13 G; 11 T; 0 other;  
Qy Query Match 32.6%; Score 57; DB 24; Length 65;  
Best Local Similarity 92.3%; Pred. No. 4.4e-10;  
Matches 60; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
Qy 26 TTCAGGAATGCACATCTCTTAAGCACTCGCAAAACAGGAGGCTCCACCTCTGGCAGG 85  
Db 1 TTCAGGAATGCACATCTCTTAAGCACTCGCAAAACAGGAGGCTCCACCTCTTAACAGC 60  
Qy 86 CCAGG 90  
Db 61 CAGGG 65  
RESULT 5  
AAK54634  
ID AAK54634 standard; cDNA; 463 BP.  
XX  
XX AAK54634;  
AC  
AC  
DT 13-NOV-2001 (first entry)  
XX  
XX Human haematological malignancy-related antigen coding sequence #359.  
DE Human; cytostatic; vascular; gene therapy; vaccine; lymphoma;  
KW haematological malignancy; antigen; chronic lymphocytic leukaemia;  
KW follicular lymphoma; Hodgkin's lymphoma; non-Hodgkin's lymphoma; ss.  
XX  
XX Homo sapiens.  
OS  
OS WO200164886-A2.  
FN  
FN 07-SEP-2001.  
XX  
XX 01-MAR-2001; 2001WO-US07272.  
PF

XX 01-MAR-2000; 2000US-0186126.  
 PR 17-MAR-2000; 2000US-0190479.  
 PR 28-APR-2000; 2000US-0200545.  
 PR 28-APR-2000; 2000US-0200303.  
 PR 28-APR-2000; 2000US-0200779.  
 PR 01-MAY-2000; 2000US-0200999.  
 PR 04-MAY-2000; 2000US-0202084.  
 PR 22-MAY-2000; 2000US-0206201.  
 PR 14-JUL-2000; 2000US-0218950.  
 PR 03-AUG-2000; 2000US-0222903.  
 PR 04-AUG-2000; 2000US-0223416.  
 PR 07-AUG-2000; 2000US-0223378.  
 XX (CORI-) CORIXA CORP.  
 XX  
 PI Gaiger A, Algate PA, Mannion J;  
 XX WPI; 2001-514842/56.  
 DR  
 XX Compositions and methods for the detection of hematological  
 PT malignancies, e.g. chronic lymphocytic leukemia, lymphoma, follicular  
 PT lymphoma and Hodgkin's and T/B cell non-Hodgkin's lymphoma.  
 XX  
 PS Claim 31; Page 426; 1252pp; English.  
 XX  
 CC The present invention relates to compositions and methods for the  
 CC detection, diagnosis and therapy of hematological malignancies. The  
 CC present sequence is the coding sequence of a human haematological  
 CC malignancy related antigen. The methods of the present invention comprise  
 CC detecting the presence of haematological malignancy related antigen(s) in  
 CC a sample obtained from the patient (an increased level of the  
 CC polypeptide, compared to an unaffected individual, is indicative of an  
 CC increased risk). Haematological malignancies which can be treated using  
 CC the present invention are chronic lymphocytic leukaemia, lymphoma,  
 CC follicular lymphoma, Hodgkin's lymphoma, non-Hodgkin's lymphoma and T/B  
 CC cell non-Hodgkin's lymphoma.  
 XX  
 SQ Sequence 463 BP; 104 A; 117 C; 88 G; 146 T; 8 other;  
 Query Match 17.6%; Score 30.8; DB 22; Length 463;  
 Best Local Similarity 53.3%; Pred. No. 1.3;  
 Matches 65; Conservative 0; Mismatches 57; Indels 0; Gaps 0;  
 QY 1 AGTCCAGGAACTTGAGCTTTGTATTTTCAGGAATGCACATCTCTTAAGCACTCGCAAAAC 60  
 DB IIIII IIIII III IIIII IIIII III IIIII III IIIII IIIII  
 265 AGTCTTCACTTCTGTTTCTTCTGCTTAACAGAGGTGCAAAATGAATTACCAACTGGGTAAAC 324  
 QY 61 AGCAAGGCTCCACACCTCTGGCAGCGCCGCTTCTCTTCAGCATGAGAAGACAGG 120  
 DB IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII  
 325 TTGAATGAGTTAATCAAAATTCAGGCCATCGTCCAGCCACCAACAGTCAAAAGGCCAAGG 384  
 QY 121 GA 122  
 DB 385 AA 386  
 RESULT 6  
 ABL11587/c  
 ID ABL11587 standard; cDNA; 2155 BP.  
 XX  
 AC ABL11587;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 29243.  
 XX  
 KW Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical; gene; ss.  
 OS Drosophila melanogaster.  
 XX  
 PN WO200171042-A2.

XX 27-SEP-2001.  
 PD  
 XX 23-MAR-2001; 2001WO-US09231.  
 PF  
 XX 23-MAR-2000; 2000US-191637P.  
 PR 11-JUL-2000; 2000US-0614150.  
 XX (PEKE ) PE CORP NY.  
 PA  
 XX Venter JC, Adams M, Li PWD, Myers EW;  
 PI  
 XX WPI; 2001-656860/75.  
 DR P-PSDB; ABB67484.  
 XX  
 PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions.  
 XX  
 PS Claim 1; SEQ ID NO 29243; 21pp + Sequence Listing; English.  
 XX  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins  
 CC (ABB57737-ABB72072).  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 2155 BP; 528 A; 607 C; 628 G; 392 T; 0 other;  
 Query Match 17.4%; Score 30.4; DB 23; Length 2155;  
 Best Local Similarity 59.1%; Pred. No. 3.5;  
 Matches 52; Conservative 0; Mismatches 36; Indels 0; Gaps 0;  
 QY 84 GCCCAGGGCGCTTCTCTTCAGCATGAGAAGAGGACAGGAGTACTCTCTCTCGG 143  
 DB IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII  
 985 GACCAAGAACACGCTCTCGGCTCGATGAGGAATATGGCAGGATTAATCGCTCTCC 926  
 QY 144 AGGACTAGTCTAGCGTAGAATAAACACC 171  
 DB IIIII IIIII IIIII IIIII IIIII IIIII IIIII IIIII  
 925 GGGACTCAGTTTGGCGCTGTTAAACTCC 898  
 RESULT 7  
 ABL11586/c  
 ID ABL11586 standard; cDNA; 14091 BP.  
 XX  
 AC ABL11586;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 29240.  
 XX  
 KW Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical; gene; ss.  
 OS Drosophila melanogaster.  
 XX  
 PN WO200171042-A2.  
 XX  
 PD 27-SEP-2001.  
 XX  
 PF 23-MAR-2001; 2001WO-US09231.  
 XX  
 PR 23-MAR-2000; 2000US-191637P.  
 PR 11-JUL-2000; 2000US-0614150.  
 XX (PEKE ) PE CORP NY.  
 PA



PS Claim 9; Page 561; 1046pp; English.

XX The present sequence is one of 3351 sequences in a library of human  
CC polynucleotides. The library is used to detect differentially expressed  
CC genes correlated with a cancerous state of a mammalian cell and can  
CC detect colon, prostate, breast and lung cancer. The library can be used  
CC to produce probes for detection of mRNA and to produce additional copies  
CC of the polynucleotides. The probes can be used for chromosome mapping of  
CC the polynucleotide and for detection of transcription levels. Ribozymes  
CC or antisense oligonucleotides can be generated. The polynucleotides and  
CC their gene products are used as genetic or biochemical markers (e.g. in  
CC blood or tissues) that will detect the earliest changes along the  
CC carcinogenesis pathway and/or monitor the efficacy of therapies and  
CC preventive interventions. The polynucleotides, polypeptides and  
CC antibodies against them can be used in pharmaceutical compositions to  
CC treat the cancers and proliferative disorders such as neoplasia,  
CC dysplasia and hyperplasia.

XX Sequence 422 BP; 89 A; 126 C; 122 G; 85 T; 0 other;

Query Match 17.1%; Score 30; DB 22; Length 422;  
Best Local Similarity 61.5%; Pred. No. 2.4;  
Matches 48; Conservative 0; Mismatches 30; Indels 0; Gaps 0;  
QY 43 CTTAGCACTCCCAAAACAGGAGGCTCCACACCTCTGGCAGGCCGCTTCTCTTC 102  
DB 210 CTGCGGTAGGAGAGACTGGAATCGGTGGCACCTCTGGGGGCGAGGCCCTTCTCTTG 151  
QY 103 AGCATGAGAAAGCAAGG 120  
DB 150 GTCAGGGGAAAGTCCAGG 133

RESULT 10

ABA42323  
ID ABA42323 standard; DNA; 483 BP.

XX ABA42323;

XX 01-FEB-2002 (first entry)

XX Human breast cell single exon nucleic acid probe #1018.

XX Human; microarray; single exon probe; gene expression; breast;  
XX disease; cancer; ss.

XX Homo sapiens.

XX WO200157271-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00662.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-496933/54.

XX New spatially-addressable set of single exon nucleic acid probes,  
XX useful for measuring gene expression in sample derived from human  
XX breast, comprises number of single exon nucleic acid probes -

PS Claim 1; SEQ ID NO 1018; 327pp + sequence listing; English.

XX

CC The invention relates to a spatially-addressable set of single exon  
CC nucleic acid probes for measuring gene expression in a sample derived  
CC from human breast and BT 474 cells. The method involves contacting  
CC the probes with a collection of detectably labelled nucleic acids  
CC derived from mRNA of human breast, and then measuring the label  
CC bound to each probe of the microarray. The probes are useful for  
CC verifying the expression of regions of genomic DNA predicted to  
CC encode proteins. They are useful for gene discovery, and for  
CC determining predisposition and/or prognosing breast disease. Gene  
CC expression analysis is useful for assessing the toxicity of chemical  
CC agents on cells. The microarray of this invention presents a far greater  
CC diversity of probes for measuring gene expression, with far less bias  
CC than expressed sequence tag microarrays. The method is suitable for  
CC rapid production of functional information from genomic sequence. The  
CC present sequence is a single exon nucleic acid probe of the invention.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 483 BP; 94 A; 136 C; 146 G; 107 T; 0 other;

Query Match 47.1%; Score 30; DB 22; Length 483;  
Best Local Similarity 61.5%; Pred. No. 2.6;  
Matches 48; Conservative 0; Mismatches 30; Indels 0; Gaps 0;  
QY 43 CTTAGCACTCCCAAAACAGGAGGCTCCACACCTCTGGCAGGCCGCTTCTCTTC 102  
DB 239 CTGCGGTAGGAGAGACTGGAATCGGTGGCACCTCTGGGGGCGAGGCCCTTCTCTTG 298  
QY 103 AGCATGAGAAAGCAAGG 120  
DB 299 GTCAGGGGAAAGTCCAGG 316

RESULT 11

ABA52750  
ID ABA52750 standard; DNA; 483 BP.

XX ABA52750;

XX 01-FEB-2002 (first entry)

XX Human foetal liver single exon nucleic acid probe #1055.

XX Human; foetal liver; gene expression; single exon nucleic acid probe; ss.

XX Homo sapiens.

XX WO200157277-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00669.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483447/52.

XX Human genome-derived single exon nucleic acid probes useful for  
XX analyzing gene expression in human fetal liver -

PS Claim 1; SEQ ID NO 1055; 639pp + sequence listing; English.

CC monitoring and prognosing diseases of the human heart and vascular system  
CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and  
CC congenital heart disease.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX Sequence 483 BP; 94 A; 136 C; 146 G; 107 T; 0 other;  
SQ

Query Match 17.1%; Score 30; DB 22; Length 483;  
Best Local Similarity 61.5%; Pred. No. 2.6;  
Matches 48; Conservative 0; Mismatches 30; Indels 0; Gaps 0;  
XX

QY 43 CTTAAGCACTCGCAAAACAGGAGGCTCCACACTCTGCGAGGCCAGGCGCTTTCCTTC 102  
Db 239 CTGCGGTAGGCAGACAGCTGGAATCGTGGCACTCTCGGGGCGAGGCCCTTCTCTT 298  
QY 103 AGCATGAGAAAGACAAGG 120  
Db 299 GTCAGGGGAAAGTCCAGG 316

RESULT 13  
AAK01005  
ID AAK01005 standard; DNA; 483 BP.  
XX  
AC AAK01005;  
XX  
DT 05-NOV-2001 (first entry)  
XX  
DE Human brain expressed single exon probe SEQ ID NO: 996.  
KW Human; brain expressed exon; gene expression analysis; probe;  
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;  
KW epilepsy; cancer; ss.  
XX  
OS Homo sapiens.  
XX  
WO WO200157275-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US00667.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
XX Penn SG, Hanzel DK, Chen W, Rank DR;  
PI WPI; 2001-483446/52.  
XX  
XX Single exon nucleic acid probes for analyzing gene expression in human  
PT brains -  
PT  
XX  
XX Example 4; SEQ ID NO: 996; 650pp + Sequence Listing; English.  
PS  
XX  
XX The present invention provides a number of single exon nucleic acid  
CC probes which are derived from genomic sequences expressed in the human  
CC brain. They can be used to measure gene expression in brain cell samples,  
CC which may enable the diagnosis and improved treatment of nervous system  
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,  
CC epilepsy and cancers. The present sequence is one of the probes of the  
CC invention.  
XX  
XX Sequence 483 BP; 94 A; 136 C; 146 G; 107 T; 0 other;  
SQ

Query Match 17.1%; Score 30; DB 22; Length 483;  
Best Local Similarity 61.5%; Pred. No. 2.6;  
Matches 48; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 43 CTTAGCACTCCGAAACAGAGGCTCCACACCTCTGCGAGGCCGCTTCTCTTC 102  
DB 239 CTGCGGTAGGAGAGACTGGAATCGGTGGCACCCTCTGCGGGCGAGGCCCTCTCTTC 298  
QY 103 AGCATGAGAAAGCAAGG 120  
DB 299 GTCAGGGGAAGTCCAGG 316

## RESULT 14

AAK26462  
ID AAK26462 standard; DNA; 483 BP.

XX AAK26462;

DT 06-NOV-2001 (first entry)

XX Human bone marrow expressed single exon probe SEQ ID NO: 1019.

XX Human; bone marrow expressed exon; gene expression analysis; probe;  
KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.

OS Homo sapiens.

PN WO200157276-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00658.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488900/53.

XX Human genome-derived single exon nucleic acid probes useful for

XX analyzing gene expression in human bone marrow -

XX Example 4; SEQ ID NO: 1019; 658pp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid  
XX probes which are derived from genomic sequences expressed in the human  
XX bone marrow. They can be used to measure gene expression in bone marrow  
XX samples, which may enable the improved diagnosis and treatment of cancers  
XX such as lymphoma, leukaemia and myeloma. The present sequence is one of  
XX the probes of the invention.

XX Sequence 483 BP; 94 A; 136 C; 146 G; 107 T; 0 other;

Query Match 17.1%; Score 30; DB 22; Length 483;  
Best Local Similarity 61.5%; Pred. No. 2.6;  
Matches 48; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 43 CTTAGCACTCCGAAACAGAGGCTCCACACCTCTGCGAGGCCGCTTCTCTTC 102  
DB 239 CTGCGGTAGGAGAGACTGGAATCGGTGGCACCCTCTGCGGGCGAGGCCCTCTCTTC 298

QY 103 AGCATGAGAAAGCAAGG 120

DB 299 GTCAGGGGAAGTCCAGG 316

## RESULT 15

AA111091

ID AA111091 standard; DNA; 483 BP.

XX AA111091;

XX 12-OCT-2001 (first entry)

DE Probe #1024 for gene expression analysis in human cervical cell sample.

XX Probe; human; microarray; gene expression; cervical epithelial cell;  
KW cervical cancer; ss.

OS Homo sapiens.

XX WO200157278-A2.

PN 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00670.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488900/53.

XX Human genome-derived single exon nucleic acid probes useful for

XX analyzing gene expression in human cervical epithelial cells -

XX Claim 25; SEQ ID No 1024; 487pp; English.

XX The present invention relates to human single exon nucleic acid probes  
XX (SENP). The present sequence is one such probe. The SENPs are derived  
XX from human HeLa cells. The SENPs can be used to produce a single exon  
XX microarray, which can be used for measuring human gene expression in a  
XX sample derived from human cervical epithelial cells. By measuring gene  
XX expression, the probes are therefore useful in grading and/or staging  
XX of diseases of the cervix, notably cervical cancer.

XX Note: The sequence data for this patent did not form part of the printed  
XX specification, but was obtained in electronic format directly from WIPO  
XX at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 483 BP; 94 A; 136 C; 146 G; 107 T; 0 other;

Query Match 17.1%; Score 30; DB 22; Length 483;  
Best Local Similarity 61.5%; Pred. No. 2.6;  
Matches 48; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 43 CTTAGCACTCCGAAACAGAGGCTCCACACCTCTGCGAGGCCGCTTCTCTTC 102  
DB 239 CTGCGGTAGGAGAGACTGGAATCGGTGGCACCCTCTGCGGGCGAGGCCCTCTCTTC 298

QY 103 AGCATGAGAAAGCAAGG 120

DB 299 GTCAGGGGAAGTCCAGG 316

Search completed: July 23, 2003, 11:15:19  
Job time : 48.3409 secs





```

US-09-182-145-18/c
; Sequence 18, Application US/09182145B
; Patent No. 6387657
; GENERAL INFORMATION:
; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDE
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/09/182
; CURRENT FILING DATE: 1998-10-29
; EARLIER APPLICATION NUMBER: US 60/066
; EARLIER FILING DATE: 1997-10-29
; EARLIER APPLICATION NUMBER: US 60/071
; EARLIER FILING DATE: 1998-02-04
; EARLIER APPLICATION NUMBER: US 60/081
; EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 18
; LENGTH: 1734
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-182-145-18

```

| Query Match           | 72.1%          | Score 126.2  | DB 4     | Length 1734 |
|-----------------------|----------------|--|----------|-------------|
| Best Local Similarity | 86.3%          | Pred. No. 1.6e-32  |          |             |
| Matches 151           | Conservative 0 | Mismatches 23  | Indels 1 | Gaps 1      |
| QY                    | 1              | AGTCCAGGAACTTGAGCTTTGTATTTTCAGGAATGCACATCTCTTAAGCACTCCGAAAC  | 60       |             |
| DB                    | 199            | AGTCCAGGAACTTGAGCTTTGTATTTCTGAATAATACACATCTCTTAATGCTCAAAAGC  | 140      |             |
| QY                    | 61             | AGGAAGGCTCCACACCTCTGGCAGGCCAGGGCCCTTCTCTTCAGCATGAGAAAGACAAAG | 120      |             |
| DB                    | 139            | AAG-AGGCTCCACATCTCTGGCAGGCCAGGGCCCTTCTCTTCAGCATGAGAGACAAAG   | 81       |             |
| QY                    | 121            | GACAGCAGATCTCTCTCTGGAGGACTAGTCTAGCCTAGAATAACACCCAAA          | 175      |             |
| DB                    | 80             | AACAGTAGATACCTCTCTCTGGAGGACTGGCCGGTCTGGATAAACACCCAAA         | 26       |             |

```

1  RESULT 3
2  US-08-652-265-1
3  ; Sequence 1, Application US/08652265
4  ; Patent No. 6025130
5  ; GENERAL INFORMATION:
6  ; APPLICANT: Thomas, Winston J.
7  ; APPLICANT: Drayna, Dennis T.
8  ; APPLICANT: Feder, John N.
9  ; APPLICANT: Gnirke, Andreas
10 ; APPLICANT: Ruddy, David
11 ; APPLICANT: Tsuchihashi, Zenta
12 ; APPLICANT: Wolff, Roger K.
13 ; TITLE OF INVENTION: Hereditary
14 ; NUMBER OF SEQUENCES: 44
15 ; CORRESPONDENCE ADDRESS:
16 ; ADDRESSEE: Townsend and Townsend
17 ; STREET: Two Embarcadero Center
18 ; CITY: San Francisco
19 ; STATE: California
20 ; COUNTRY: USA
21 ; ZIP: 94111-3834
22 ; COMPUTER READABLE FORM:
23 ; MEDIUM TYPE: Floppy disk
24 ; COMPUTER: IBM PC compatible
25 ; OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/652,265
FILING DATE: 23-MAY-1996
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Smith, William M.
REGISTRATION NUMBER: 30,223
REFERENCE/DOCKET NUMBER: 17957-000500
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 10825 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: join(361..436, 3762..4025, 4235..4510, 5606..5881,
LOCATION: 6040..6153, 7107..7147)
OTHER INFORMATION: {/product="Hereditary Hemochromatosis
OTHER INFORMATION:
OTHER INFORMATION: {/note="No. 6025130mal or wild-type (unaffected)
OTHER INFORMATION: {/hereditary Hemochromatosis (HH) gene
OTHER INFORMATION: {/allele"
FEATURE:

```

```

1 NAME/NCBI: LOCATION: 140..7319 /note= "start and stop positions for
2 OTHER INFORMATION: normal or wild-type (unaffected) allele
3 OTHER INFORMATION: CDNA (SEQ ID NO:3)"
4 FEATURE:
5 NAME/KEY: -
6 LOCATION: 3852..3893 /note= "start and stop positions for
7 OTHER INFORMATION: normal or wild-type (unaffected) genomic
8 OTHER INFORMATION: sequence surrounding variant for 24d2(c)
9 OTHER INFORMATION: allele (SEQ ID NO:41)"
10 FEATURE:

```

```
,
,
, NAME/KEY: -
, LOCATION: 5507..6023
, OTHER INFORMATION: /note= "start and stop positions for
, OTHER INFORMATION: normal or wild-type (unaffected) genomic
, OTHER INFORMATION: sequence surrounding variant for 24d1(G)
, OTHER INFORMATION: allele (SEQ ID NO:20)"
,
```

```

1 NAME/KEY: allele
2 LOCATION: replace(3872, "c")
3 OTHER INFORMATION: /phenotype= "normal or wild-type
4 OTHER INFORMATION: /unaffected="
5 OTHER INFORMATION: /label= 24d2
6
7 FEATURE:
8 NAME/KEY: allele
9 LOCATION: replace(3898, "a")
10 OTHER INFORMATION: /phenotype= "normal or wild-type
11 OTHER INFORMATION: /unaffected="
12 OTHER INFORMATION: /label= 24d7
13
14 FEATURE:
15 NAME/KEY: allele
16 LOCATION: replace(5834, "g")
17 OTHER INFORMATION: /phenotype= "normal or wild-type
18 OTHER INFORMATION: /unaffected="
19 OTHER INFORMATION: /label= 24d1
20
21 C 00 5834 5835

```

Query Match 15.9%; Score 27.8; DB 3; Length 10825;  
Best Local Similarity 65.1%; Pred. No. 11;  
Matches 41; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

```

; NAME/KEY: 5507...6023
; LOCATION: /note= "start and stop positions for
; OTHER INFORMATION: genomic sequence surrounding variant
; OTHER INFORMATION: for 24d1(A) allele (SEQ ID NO:21)"
; FEATURE:
; NAME/KEY: allele
; LOCATION: replace(5834, "a")
; OTHER INFORMATION: /phenotype= "Hereditary Hemochromatosis
; OTHER INFORMATION:
; OTHER INFORMATION: /label= 24d1
US-08-652-265-3
Query Match 15.9%; Score 27.8; DB 3; Length 10825;
Best Local Similarity 65.1%; Pred. No. 11;
Matches 41; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
QY 113 AGCAAGGGACACAGAGTACTCTCTCTGAGAGACTAGTCTAGCCTAGATAAACACCC 172
    || || || || || || || || || || || || || || || || || || || || ||
Db 2609 AGCCACAGGAGACAGATTCTCTGAGTCTCAAGACCAAGCCTGGCGCAACACAGCA 2668
QY 173 AAA 175
    |||
Db 2669 AAA 2671

RESULT 5
US-08-652-265-5
; Sequence 5, Application US/08652265
; Patent No. 6025130
; GENERAL INFORMATION:
; APPLICANT: Thomas, Winston J.
; APPLICANT: Drayna, Dennis T.
; APPLICANT: Feder, John N.
; APPLICANT: Gnirke, Andreas
; APPLICANT: Ruddy, David
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Hereditary Hemochromatosis Gene
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/652,265
; FILING DATE: 23-MAY-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M.
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 17957-000500
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10825 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(361..436, 3762..4025, 4235..4510, 5606..5881).

```

LOCATION: 6040..6153, 7107..7147)  
OTHER INFORMATION: /product= "Hereditary Hemochromatosis"  
OTHER INFORMATION: mutation  
OTHER INFORMATION: /note= "Hereditary Hemochromatosis (HH)"  
OTHER INFORMATION: gene 24d2 allele"  
FEATURE:  
NAME/KEY: -  
LOCATION: 140..7319  
OTHER INFORMATION: /note= "start and stop positions for  
OTHER INFORMATION: 24d2 allele CDNA (SEQ ID NO:11)"  
FEATURE:  
NAME/KEY: -  
LOCATION: 3852..3891  
OTHER INFORMATION: /note= "start and stop positions for  
OTHER INFORMATION: genomic sequence surrounding variant  
OTHER INFORMATION: for 24d2(G) allele (SEQ ID NO:42)"  
FEATURE:  
NAME/KEY: -  
LOCATION: 5507..6023  
OTHER INFORMATION: /note= "start and stop positions for  
OTHER INFORMATION: genomic sequence surrounding variant  
OTHER INFORMATION: for 24d1(G) allele (SEQ ID NO:20)"  
FEATURE:  
NAME/KEY: allele  
LOCATION: replace(3872, "g")  
OTHER INFORMATION: /phenotype= "Hereditary Hemochromatosis"  
OTHER INFORMATION: /label= 24d2  
US-08-652-265-5

Query Match 15.9%; Score 27.8; DB 3; Length 10825;  
Best Local Similarity 65.1%; Pred. No. 11;  
Matches 41; Conservative 0; Mismatches 22; Indels 0; Gaps 0;  
QY 113 AGACAAGGAGCAGACGACTCTCTCTGGAGGACTAGTCTAGCTAGATAACACCC 172  
Db 2609 AGGCAAGGAGCAGACGACTCTCTCTGGAGGACTAGTCTAGCTAGATAACACCC 172  
QY 173 AAA 175  
Db 2669 AAA 2671

RESULT 6  
US-08-652-265-7  
Sequence 7, Application US/08652265  
Patent No. 6025130  
GENERAL INFORMATION:  
APPLICANT: Thomas, Winston J.  
APPLICANT: Drayna, Dennis T.  
APPLICANT: Feder, John N.  
APPLICANT: Gnirke, Andreas  
APPLICANT: Ruddy, David  
APPLICANT: Tsuchihashi, Zenta  
APPLICANT: Wolff, Roger K.  
TITLE OF INVENTION: Hereditary Hemochromatosis Gene  
NUMBER OF SEQUENCES: 44  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/652,265  
FILING DATE: 23-MAY-1996

CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, William M.  
REGISTRATION NUMBER: 30,223  
REFERENCE/DOCKET NUMBER: 17957-000500  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10825 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: join(361..436, 3762..4025, 4235..4510, 5606..5881,  
LOCATION: 6040..6153, 7107..7147)  
OTHER INFORMATION: /product= "Hereditary Hemochromatosis"  
OTHER INFORMATION:  
OTHER INFORMATION: and 24d2 mutations"  
OTHER INFORMATION: /note= "Hereditary Hemochromatosis (HH)"  
OTHER INFORMATION: gene containing a combination of both  
OTHER INFORMATION: 24d1 and 24d2 alleles"  
FEATURE:  
NAME/KEY: -  
LOCATION: 140..7319  
OTHER INFORMATION: /note= "start and stop positions for  
OTHER INFORMATION: CDNA containing a combination of both  
OTHER INFORMATION: 24d1 and 24d2 alleles  
OTHER INFORMATION: (SEQ ID NO:12)"  
FEATURE:  
NAME/KEY: -  
LOCATION: 3852..3891  
OTHER INFORMATION: /note= "start and stop positions for  
OTHER INFORMATION: genomic sequence surrounding variant  
OTHER INFORMATION: for 24d2(G) allele (SEQ ID NO:42)"  
FEATURE:  
NAME/KEY: -  
LOCATION: 5507..6023  
OTHER INFORMATION: /note= "start and stop positions for  
OTHER INFORMATION: genomic sequence surrounding variant  
OTHER INFORMATION: for 24d1(A) allele (SEQ ID NO:21)"  
FEATURE:  
NAME/KEY: allele  
LOCATION: replace(3872, "g")  
OTHER INFORMATION: /phenotype= "Hereditary Hemochromatosis"  
OTHER INFORMATION: /label= 24d2  
FEATURE:  
NAME/KEY: allele  
LOCATION: replace(5834, "a")  
OTHER INFORMATION: /phenotype= "Hereditary Hemochromatosis"  
OTHER INFORMATION: /label= 24d1  
US-08-652-265-7

Query Match 15.9%; Score 27.8; DB 3; Length 10825;  
Best Local Similarity 65.1%; Pred. No. 11;  
Matches 41; Conservative 0; Mismatches 22; Indels 0; Gaps 0;  
QY 113 AGACAAGGAGCAGACGACTCTCTCTGGAGGACTAGTCTAGCTAGATAACACCC 172  
Db 2609 AGGCAAGGAGCAGACGACTCTCTCTGGAGGACTAGTCTAGCTAGATAACACCC 172  
QY 173 AAA 175  
Db 2669 AAA 2671

RESULT 7  
US-08-834-497A-1

```

; Sequence 1, Application US/08834497A
; Patent No. 6140305
; GENERAL INFORMATION:
; APPLICANT: Thomas, Winston J.
; APPLICANT: Drayna, Dennis T.
; APPLICANT: Feder, John N.
; APPLICANT: Gnirke, Andreas
; APPLICANT: Ruddy, David
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: HEREDITARY HEMOCHROMATOSIS GENE PRODUCTS
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/834,497A
; FILING DATE: 04-APR-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/652,265
; FILING DATE: 23-MAY-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/632,673
; FILING DATE: 16-APR-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/630,912
; FILING DATE: 04-APR-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0056-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-493-4935
; TELEFAX: 650-493-5556
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10825 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(361..436, 3762..4025, 4235..4510, 5606..5881,
; LOCATION: 6040..6153, 7107..7147)
; OTHER INFORMATION: /product= "Hereditary Hemochromatosis
; OTHER INFORMATION: /note= "No. 6140305mal or wild-type (unaffected)
; OTHER INFORMATION: Hereditary Hemochromatosis (HH) gene
; OTHER INFORMATION: allele"
; FEATURE:
; NAME/KEY: -
; LOCATION: 140..7319
; OTHER INFORMATION: /note= "start and stop positions for
; OTHER INFORMATION: normal or wild-type (unaffected) allele
; OTHER INFORMATION: cDNA (SEQ ID NO:9)"
; FEATURE:
; NAME/KEY: -
; LOCATION: 3852..3891

```

```

: OTHER INFORMATION: /note= "start and stop positions for
: OTHER INFORMATION: normal or wild-type (unaffected) genomic
: OTHER INFORMATION: sequence surrounding variant for 24d2(C)
: OTHER INFORMATION: allele (SEQ ID NO:41)"
: FEATURE:
: NAME/KEY: -
: LOCATION: 5507..6023
: OTHER INFORMATION: /note= "start and stop positions for
: OTHER INFORMATION: normal or wild-type (unaffected) genomic
: OTHER INFORMATION: sequence surrounding variant for 24d1(G)
: OTHER INFORMATION: allele (SEQ ID NO:20)"
: FEATURE:
: NAME/KEY: allele
: LOCATION: replace(3872, "c")
: OTHER INFORMATION: /phenotype= "normal or wild-type
: OTHER INFORMATION: (unaffected)"
: OTHER INFORMATION: /label= 24d2
: FEATURE:
: NAME/KEY: allele
: LOCATION: replace(3878, "a")
: OTHER INFORMATION: /phenotype= "normal or wild-type
: OTHER INFORMATION: (unaffected)"
: OTHER INFORMATION: /label= 24d7
: FEATURE:
: NAME/KEY: allele
: LOCATION: replace(5834, "g")
: OTHER INFORMATION: /phenotype= "normal or wild-type
: OTHER INFORMATION: (unaffected)"
: OTHER INFORMATION: /label= 24d1
: US-08-834-497A-1
:
: Query Match 15.9%; Score 27.8; DB 3; Length 10825;
: Best Local Similarity 65.1%; Pred. No. 11;
: Matches 41; Conservative 0; Mismatches 22; Indels 0;
:
: QY 113 AGACAAGGACAGCAGAGTACTCTCCCTGTGGAGGACTAGTCTAGCCTAGATAATA
: Db 2609 AGGCCAAGGACAGCAGATTCTTGAGCTCAGGAGTTCAAGACCAGCCTGGCAAA
:
: QY 173 AAA 175
: Db 2669 AAA 2671
:
: RESULT 8
: US-08-834-497A-3
: Sequence 3, Application US/08834497A
: Patent No. 6140305
: GENERAL INFORMATION:
: APPLICANT: Thomas, Winston J.
: APPLICANT: Drayna, Dennis T.
: APPLICANT: Feder, John N.
: APPLICANT: Gnirke, Andreas
: APPLICANT: Ruddy, David
: APPLICANT: Tsuchihashi, Zenta
: APPLICANT: Wolff, Roger K.
: TITLE OF INVENTION: HEREDITARY HEMOCHROMATOSIS GENE PRODUCTS
: NUMBER OF SEQUENCES: 76
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Pennie & Edmonds LLP
: STREET: 1155 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: USA
: ZIP: 10036-2811
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: Windows 95
: SOFTWARE: FASTSEQ for Windows Version 2.0b
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/834,497A
: FILING DATE: 04-APR-1997

```



```

;
; FEATURE:
; NAME/KEY: 3852..3891
; LOCATION: join(361..436, 3762..4025, 4235..4510, 5606..5881,
; OTHER INFORMATION: /note= "start and stop positions for
; OTHER INFORMATION: genomic sequence surrounding variant
; OTHER INFORMATION: for 24d2(G) allele (SEQ ID NO:42)"
;
; FEATURE:
; NAME/KEY: 5507..6023
; LOCATION: join(361..436, 3762..4025, 4235..4510, 5606..5881,
; OTHER INFORMATION: /note= "start and stop positions for
; OTHER INFORMATION: genomic sequence surrounding variant
; OTHER INFORMATION: for 24d1(G) allele (SEQ ID NO:20)"
;
; FEATURE:
; NAME/KEY: allele
; LOCATION: replace(3872, "g")
; OTHER INFORMATION: /phenotype= "Hereditary Hemochromatosis"
; OTHER INFORMATION: /label= 24d2
;
; US-08-834-497A-5
;
; Query Match 15.9%; Score 27.8; DB 3; Length 10825;
; Best Local Similarity 65.1%; Pred. No. 11;
; Matches 41; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
;
; Qy 113 AGACAAGGGACAGCAGAGTACTCTCTCTGGAGGACTAGTCTAGCCTAGATAAACACCC 172
;   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
; Db 2609 AGGCAAGGAGAGCAGAGTCTCTGAGCTCAGGAGTTCAAGACCAGCCTGGGCAACACAGCA 2668
;
; Qy 173 AAA 175
;   |||
; Db 2669 AAA 2671
;
; RESULT 10
; US-08-834-497A-7
; Sequence 7, Application US/08834497A
; Patent No. 6140305
; GENERAL INFORMATION:
; APPLICANT: Thomas, Winston J.
; APPLICANT: Drayna, Dennis T.
; APPLICANT: Feder, John N.
; APPLICANT: Gnirke, Andreas
; APPLICANT: Ruddy, David
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: HEREDITARY HEMOCHROMATOSIS GENE PRODUCTS
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/834,497A
; FILING DATE: 04-APR-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/652,265
; FILING DATE: 23-MAY-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/632,673
; FILING DATE: 16-APR-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/630,912
;
; FILING DATE: 04-APR-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0056-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-493-4935
; TELEFAX: 650-493-5556
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10825 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(361..436, 3762..4025, 4235..4510, 5606..5881,
; OTHER INFORMATION: /product= "Hereditary Hemochromatosis"
; OTHER INFORMATION: and 24d2 mutations"
; OTHER INFORMATION: /note= "Hereditary Hemochromatosis (HH)"
; OTHER INFORMATION: gene containing a combination of both
; OTHER INFORMATION: 24d1 and 24d2 alleles"
; FEATURE:
; NAME/KEY:
; LOCATION: 140..7319
; OTHER INFORMATION: /note= "start and stop positions for
; OTHER INFORMATION: CDNA containing a combination of both
; OTHER INFORMATION: 24d1 and 24d2 alleles"
; OTHER INFORMATION: (SEQ ID NO:12)"
; FEATURE:
; NAME/KEY:
; LOCATION: 3852..3891
; OTHER INFORMATION: /note= "start and stop positions for
; OTHER INFORMATION: genomic sequence surrounding variant
; OTHER INFORMATION: for 24d2(G) allele (SEQ ID NO:42)"
; FEATURE:
; NAME/KEY:
; LOCATION: 5507..6023
; OTHER INFORMATION: /note= "start and stop positions for
; OTHER INFORMATION: genomic sequence surrounding variant
; OTHER INFORMATION: for 24d1(A) allele (SEQ ID NO:21)"
; FEATURE:
; NAME/KEY: allele
; LOCATION: replace(3872, "g")
; OTHER INFORMATION: /phenotype= "Hereditary Hemochromatosis"
; OTHER INFORMATION: /label= 24d1
;
; US-08-834-497A-7
;
; Query Match 15.9%; Score 27.8; DB 3; Length 10825;
; Best Local Similarity 65.1%; Pred. No. 11;
; Matches 41; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
;
; Qy 113 AGACAAGGGACAGCAGAGTACTCTCTCTGGAGGACTAGTCTAGCCTAGATAAACACCC 172
;   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
; Db 2609 AGGCAAGGAGAGCAGAGTCTCTGAGCTCAGGAGTTCAAGACCAGCCTGGGCAACACAGCA 2668
;
; Qy 173 AAA 175
;   |||
; Db 2669 AAA 2671
;   |||
```

```

; OTHER INFORMATION: normal or wild-type (unaffected) genomic
; OTHER INFORMATION: sequence surrounding variant for 24d2(C)
; OTHER INFORMATION: allele (SEQ ID NO:41)"
;
; NAME/KEY: 5507..6023
; LOCATION:
; OTHER INFORMATION: /note= "start and stop positions for
; OTHER INFORMATION: normal or wild-type (unaffected) genomic
; OTHER INFORMATION: sequence surrounding variant for 24d1(G)
; OTHER INFORMATION: allele (SEQ ID NO:20)"
;
; FEATURE:
; NAME/KEY: allele
; LOCATION: replace(3872, "c")
; OTHER INFORMATION: /phenotype= "normal or wild-type
; OTHER INFORMATION: (unaffected)"
; OTHER INFORMATION: /label= 24d2
;
; FEATURE:
; NAME/KEY: allele
; LOCATION: replace(3878, "a")
; OTHER INFORMATION: /phenotype= "normal or wild-type
; OTHER INFORMATION: (unaffected)"
; OTHER INFORMATION: /label= 24d7
;
; NAME/KEY: allele
; LOCATION: replace(5834, "g")
; OTHER INFORMATION: /phenotype= "normal or wild-type
; OTHER INFORMATION: (unaffected)"
; OTHER INFORMATION: /label= 24d1
;
; US-09-503-444A-1
;
; Query Match 15.9%; Score 27.8; DB 4; Length 10825;
; Best Local Similarity 65.1%; Pred. No. 11;
; Matches 41; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
;
; QY 113 AGACAAGGACAGCAGAGACTCTCTCTGGAGGACTAGTCTGACCTAGCAATAAACACCC 172
; | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
; Db 2609 AGCCCAAGGAGAGAGAGATTCTTGAGTCAGGAGTTCAAGACCAGCCTGGCGAACACAGCA 2668
;
; QY 173 AAA 175
; | | |
; Db 2669 AAA 2671
;
; RESULT 12
; US-09-503-444A-3
; Sequence 3, Application US/09503444A
; Patent No. 6228594
; GENERAL INFORMATION:
; APPLICANT: Thomas, Winston J.
; APPLICANT: Drayna, Dennis T.
; APPLICANT: Feder, John N.
; APPLICANT: Gnirke, Andreas
; APPLICANT: Ruddy, David
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Hereditary Hemochromatosis Gene
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect Version 8
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/503,444A
; FILING DATE: 14-Feb-2000
; CLASSIFICATION:

```



```

: Sequence 5, Application US/09503444A
: Patent No. 6228594
: GENERAL INFORMATION:
: APPLICANT: Thomas, Winston J.
: APPLICANT: Drayna, Dennis T.
: APPLICANT: Feder, John N.
: APPLICANT: Gnirke, Andreas
: APPLICANT: Ruddy, David
: APPLICANT: Tsuchihashi, Zenta
: APPLICANT: Wolff, Roger K.
: TITLE OF INVENTION: Hereditary Hemochromatosis Gene
: NUMBER OF SEQUENCES: 44
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Pennie & Edmonds LLP
: STREET: 1155 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: USA
: ZIP: 10036
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: Windows 95
: SOFTWARE: WordPerfect version 8
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/503,444A
: FILING DATE: 14-Feb-2000
: CLASSIFICATION:
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: 08/652,265
: FILING DATE: 23-May-1996
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: 08/632,673
: FILING DATE: 16-Apr-1996
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: 08/630,912
: FILING DATE: 04-Apr-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Poissant, Brian M.
: REGISTRATION NUMBER: 28,462
: REFERENCE/DOCKET NUMBER: 8907-0088-999
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 212-790-9090
: TELEFAX: 212-869-9741
: TELEX: 66141
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 10825 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: join(361..436, 3762..4025, 4235..4510, 5606..5881,
: LOCATION: 6040..6153, 7107..7147)
: OTHER INFORMATION: /product= "Hereditary Hemochromatosis"
: OTHER INFORMATION:
: OTHER INFORMATION: mutation"
: OTHER INFORMATION: /note= "Hereditary Hemochromatosis (HH)
: OTHER INFORMATION: gene 24d2 allele"
: FEATURE:
: NAME/KEY: -
: LOCATION: 140..7319
: OTHER INFORMATION: /note= "start and stop positions for
: OTHER INFORMATION: 24d2 allele cDNA (SEQ ID NO:11)"
: FEATURE:
: NAME/KEY: -
: LOCATION: 3852..3891
: OTHER INFORMATION: /note= "start and stop positions for
: OTHER INFORMATION: genomic sequence surrounding variant
: OTHER INFORMATION: for 24d2(c) allele (SEQ ID NO:42)"
: FEATURE:

```

```
; NAME/KEY: 5507..6023
; LOCATION: join(361..436, 3762..4025, 4235..4510, 5606..5881,
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10825 base pairs
; TYPE: nucleic acid
; OTHER INFORMATION: genomic sequence surrounding variant
; OTHER INFORMATION: for 24d1(G) allele (SEQ ID NO:20)"
; FEATURE:
; NAME/KEY: allele
; LOCATION: replace(3872, "g")
; OTHER INFORMATION: /phenotype= "Hereditary Hemochromatosis
; OTHER INFORMATION: /label= 24d2
; OTHER INFORMATION:
; US-09-503-444A-5

Query Match 15.9%; Score 27.8; DB 4; Length 10825;
Best Local Similarity 65.1%; Pred. No. 11;
Matches 41; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 113 AGACAAGGACAGCAGACTCTCTCTGGAGGACTAGTCTAGCTAGATAAACACCC 172
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2609 AGGCAAGGACAGCAGACTCTCTGAGCTCAGGAGTTCAAGACCAGCCTGGCAACACGCA 2668

QY 173 AAA 175
|||
Db 2669 AAA 2671

RESULT 14
US-09-503-444A-7
; Sequence 7, Application US/09503444A
; Patent No. 6228594
; GENERAL INFORMATION:
; APPLICANT: Thomas, Winston J.
; APPLICANT: Drayna, Dennis T.
; APPLICANT: Feder, John N.
; APPLICANT: Gairke, Andreas
; APPLICANT: Ruddy, David
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Hereditary Hemochromatosis Gene
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect Version 8
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/503,444A
; FILING DATE: 14-Feb-2000
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/652,265
; FILING DATE: 23-May-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/632,673
; FILING DATE: 16-Apr-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/630,912
; FILING DATE: 04-Apr-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0088-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-9741
; TELEX: 66141
```

```
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10825 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(361..436, 3762..4025, 4235..4510, 5606..5881,
; LOCATION: 6040..6153, 7107..7147)
; OTHER INFORMATION: /product= "Hereditary Hemochromatosis
; OTHER INFORMATION:
; OTHER INFORMATION: and 24d2 mutations"
; OTHER INFORMATION: /note= "Hereditary Hemochromatosis (HH)
; OTHER INFORMATION: gene containing a combination of both
; OTHER INFORMATION: 24d1 and 24d2 alleles"
; FEATURE:
; NAME/KEY:
; LOCATION: 140..7319
; OTHER INFORMATION: /note= "start and stop positions for
; OTHER INFORMATION: cDNA containing a combination of both
; OTHER INFORMATION: 24d1 and 24d2 alleles
; OTHER INFORMATION: {SEQ ID NO:12}"
; FEATURE:
; NAME/KEY:
; LOCATION: 3852..3891
; OTHER INFORMATION: /note= "start and stop positions for
; OTHER INFORMATION: genomic sequence surrounding variant
; OTHER INFORMATION: for 24d2(G) allele (SEQ ID NO:42)"
; FEATURE:
; NAME/KEY:
; LOCATION: 5507..6023
; OTHER INFORMATION: /note= "start and stop positions for
; OTHER INFORMATION: genomic sequence surrounding variant
; OTHER INFORMATION: for 24d1(A) allele (SEQ ID NO:21)"
; FEATURE:
; NAME/KEY: allele
; LOCATION: replace(3872, "g")
; OTHER INFORMATION: /phenotype= "Hereditary Hemochromatosis
; OTHER INFORMATION:
; OTHER INFORMATION: /label= 24d2
; FEATURE:
; NAME/KEY: allele
; LOCATION: replace(5894, "a")
; OTHER INFORMATION: /phenotype= "Hereditary Hemochromatosis
; OTHER INFORMATION:
; OTHER INFORMATION: /label= 24d1
; US-09-503-444A-7

Query Match 15.9%; Score 27.8; DB 4; Length 10825;
Best Local Similarity 65.1%; Pred. No. 11;
Matches 41; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 113 AGACAAGGACAGCAGACTCTCTCTGGAGGACTAGTCTAGCTAGATAAACACCC 172
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2609 AGGCAAGGACAGCAGACTCTCTGAGCTCAGGAGTTCAAGACCAGCCTGGCAACACGCA 2668

QY 173 AAA 175
|||
Db 2669 AAA 2671

RESULT 15
US-09-277-457-27
; Sequence 27, Application US/09277457
; Patent No. 6355425
; GENERAL INFORMATION:
; APPLICANT: Rothenberg, Barry E.
; APPLICANT: Sawada-Hirai, Ritsuko
; APPLICANT: Barton, James G.
; TITLE OF INVENTION: MUTATIONS ASSOCIATED WITH IRON DISORDERS
; FILE REFERENCE: 10653/002001
```





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 15:36:41 ; Search time 43.1391 Seconds  
(without alignments)  
8368.853 Million cell updates/sec

Title: US-10-010-408-1\_COPY\_1534\_1708

Perfect score: 175

Sequence: 1 AGTCCAGGAAGCTTGAGCTTT.....GCCTAGATAAACACCCAAA 175

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1439767 seqs, 1031500376 residues

Total number of hits satisfying chosen parameters: 2879534

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published\_Applications\_NA.\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq1.\*  
11: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq2.\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq3.\*  
13: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
15: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description         |
|------------|-------|-------------|--------|----|---------------------|
| 1          | 175   | 100.0       | 1708   | 15 | US-10-010-408-1     |
| 2          | 126.2 | 72.1        | 1734   | 15 | Sequence 1, Appl    |
| 3          | 126.2 | 72.1        | 1734   | 15 | Sequence 17, Appl   |
| 4          | 125.8 | 71.9        | 439    | 12 | US-10-112-267-17    |
| 5          | 31.2  | 17.8        | 707    | 15 | US-10-112-267-18    |
| 6          | 30.8  | 17.6        | 463    | 11 | US-09-956-622A-23   |
| 7          | 30.8  | 17.6        | 463    | 11 | US-10-027-632-12045 |
| 8          | 30.8  | 17.6        | 463    | 11 | Sequence 359, App   |
| 9          | 30.8  | 17.6        | 463    | 11 | Sequence 4928, App  |
| 10         | 30    | 17.1        | 483    | 10 | US-09-796-692-359   |
| 11         | 30    | 17.1        | 1096   | 10 | US-10-040-862-359   |
| 12         | 29.4  | 16.8        | 624    | 15 | US-10-040-862-4928  |
| 13         | 29    | 16.6        | 525    | 15 | US-09-864-761-1000  |
| 14         | 29    | 16.6        | 851    | 15 | US-09-864-761-1000  |
| 15         | 28.8  | 16.5        | 431    | 11 | Sequence 17772, A   |
| 16         | 28.8  | 16.5        | 659158 | 11 | Sequence 23, Appl   |
|            |       |             |        |    | Sequence 205550,    |
|            |       |             |        |    | Sequence 110489,    |
|            |       |             |        |    | Sequence 222, App   |
|            |       |             |        |    | Sequence 20, Appl   |

#### ALIGNMENTS

##### RESULT 1

US-10-010-408-1  
; Sequence 1, Application US/10010408  
; Publication No. US20020165185A1  
; GENERAL INFORMATION:

APPLICANT: John J. Castellet, Jr.

TITLE OF INVENTION: No. US20020165185A1el Heparin-Induced CCN-Like Molecules and Uses Therefor

NUMBER OF SEQUENCES: 13

CORRESPONDENCE ADDRESS:

ADDRESSEE: LAHIVE & COCKFIELD, LLP

STREET: 28 State Street

CITY: Boston

STATE: Massachusetts

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/010,408

FILING DATE: 07-Dec-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/044,273

FILING DATE: March 19, 1998

APPLICATION NUMBER: <Unknown>

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Amy E. Mandragouras

REGISTRATION NUMBER: 36,207

REFERENCE/DOCKET NUMBER: MBI-004

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617)227-7400

TELEFAX: (617)742-4214

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

Sequence 281405,  
Sequence 189371,  
Sequence 189372,  
Sequence 3889, Ap  
Sequence 6, Appl  
Sequence 12168, A  
Sequence 1852, Ap  
Sequence 9, Appl  
Sequence 174961,  
Sequence 12167, A  
Sequence 3909, Ap  
Sequence 21981, A  
Sequence 21982, A  
Sequence 101680,  
Sequence 9766, Ap  
Sequence 9765, Ap  
Sequence 296356,  
Sequence 5700, Ap  
Sequence 134, App  
Sequence 25, Appl  
Sequence 27, Appl  
Sequence 3327, Ap  
Sequence 1, Appl  
Sequence 2, Appl  
Sequence 29616, A  
Sequence 131, App  
Sequence 9209, Ap  
Sequence 284014,

LENGTH: 1708 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 249..1001  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-10-010-408-1

Query Match 100.0%; Score 175; DB 15; Length 1708;  
Best Local Similarity 100.0%; Pred. No. 7.2e-53;  
Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AGTCAGGAGTGTGAGCTTTGTTATTTTCAGGAATGCACATCTCTTAAGCACTCGCAAAAC 60  
DB 1534 AGTCAGGAGTGTGAGCTTTGTTATTTTCAGGAATGCACATCTCTTAAGCACTCGCAAAAC 1593  
QY 61 AGGAAGGCTCCACACTCTGGCAGGCCAGGGCTTTCTTTCAGCATGAGAAGACAAGG 120  
DB 1594 AGGAAGGCTCCACACTCTGGCAGGCCAGGGCTTTCTTTCAGCATGAGAAGACAAGG 1653  
QY 121 GACAGCAGAGTACTCTCTCTCGGAGCACTAGTCTAGCCCTAGATTAACACCCAAA 175  
DB 1654 GACAGCAGAGTACTCTCTCTCGGAGCACTAGTCTAGCCCTAGATTAACACCCAAA 1708

## RESULT 2

US-10-112-267-17  
Sequence 17, Application US/10112267  
Publication No. US20030068678A1

## GENERAL INFORMATION:

APPLICANT: Botstein, David A.  
APPLICANT: Cohen, Robert  
APPLICANT: Goddard, Audrey  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: Lawrence, David A.  
APPLICANT: Levine, Arnold J.  
APPLICANT: Pennica, Diane  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME

FILE REFERENCE: P1176R2

CURRENT APPLICATION NUMBER: US/10/112,267

PRIOR FILING DATE: 2002-03-27

PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/182,145B

PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-29

PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/063,704

PRIOR FILING DATE: EARLIER FILING DATE: 1997-10-29

PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/073,612

PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-04

PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/081,695

PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-14

NUMBER OF SEQ ID NOS: 156

SEQ ID NO 17

LENGTH: 1734

TYPE: DNA

ORGANISM: Mus musculus

US-10-112-267-17

Query Match 72.1%; Score 126.2; DB 15; Length 1734;  
Best Local Similarity 86.3%; Pred. No. 3.1e-35;  
Matches 151; Conservative 0; Mismatches 23; Indels 1; Gaps 1;

QY 1 AGTCAGGAGTGTGAGCTTTGTTATTTTCAGGAATGCACATCTCTTAAGCACTCGCAAAAC 60

DB 1536 AGTCAGGAGTGTGAGCTTTGTTATTTTCAGGAATGCACATCTCTTAAGCACTCGCAAAAC 1595

QY 61 AGGAAGGCTCCACACTCTGGCAGGCCAGGGCTTTCTTTCAGCATGAGAAGACAAGG 120

DB 1596 AAG-AGGCTCCACACTCTGGCAGGCCAGGGCTTTCTTTCAGCATGAGAAGACAAGG 1654

QY 121 GACAGCAGAGTACTCTCTCTCGGAGCACTAGTCTAGCCCTAGATTAACACCCAAA 175  
DB 1655 AACAGTAGAGTACCTCTCTCTCGGAGCACTAGTCTAGCCCTAGATTAACACCCAAA 1709

## RESULT 3

US-10-112-267-18/c  
Sequence 18, Application US/10112267  
Publication No. US20030068678A1

## GENERAL INFORMATION:

APPLICANT: Botstein, David A.  
APPLICANT: Cohen, Robert  
APPLICANT: Goddard, Audrey  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: Lawrence, David A.  
APPLICANT: Levine, Arnold J.  
APPLICANT: Pennica, Diane  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME  
FILE REFERENCE: P1176R2  
CURRENT APPLICATION NUMBER: US/10/112,267

CURRENT FILING DATE: 2002-03-27

PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/182,145B

PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-29

PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/063,704

PRIOR FILING DATE: EARLIER FILING DATE: 1997-10-29

PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/073,612

PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-04

PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/081,695

PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-14

NUMBER OF SEQ ID NOS: 156

SEQ ID NO 18

LENGTH: 1734

TYPE: DNA

ORGANISM: Mus musculus

US-10-112-267-18

Query Match 72.1%; Score 126.2; DB 15; Length 1734;  
Best Local Similarity 86.3%; Pred. No. 3.1e-35;  
Matches 151; Conservative 0; Mismatches 23; Indels 1; Gaps 1;

QY 1 AGTCAGGAGTGTGAGCTTTGTTATTTTCAGGAATGCACATCTCTTAAGCACTCGCAAAAC 60

DB 199 AGTCAGGAGTGTGAGCTTTGTTATTTTCAGGAATGCACATCTCTTAAGCACTCGCAAAAC 140

QY 61 AGGAAGGCTCCACACTCTGGCAGGCCAGGGCTTTCTTTCAGCATGAGAAGACAAGG 120

DB 139 AAG-AGGCTCCACACTCTGGCAGGCCAGGGCTTTCTTTCAGCATGAGAAGACAAGG 81

QY 121 GACAGCAGAGTACTCTCTCTCGGAGCACTAGTCTAGCCCTAGATTAACACCCAAA 175

DB 80 AACAGTAGAGTACTCTCTCTCGGAGCACTAGTCTAGCCCTAGATTAACACCCAAA 26

## RESULT 4

US-09-956-622A-23  
Sequence 23, Application US/09956622A  
Publication No. US20030091973A1

## GENERAL INFORMATION:

APPLICANT: Horesovsky, Gregory J  
APPLICANT: No. US20030091973A1 II, L. Staton  
APPLICANT: Raha, Debashish  
TITLE OF INVENTION: Method of Identifying Osteoregenerative Agents Using  
FILE OF INVENTION: Differential Gene Expression  
FILE REFERENCE: 21402-445

CURRENT APPLICATION NUMBER: US/09/956,622A

CURRENT FILING DATE: 2001-09-19

PRIOR APPLICATION NUMBER: 60/233,579

PRIOR FILING DATE: 2000-09-19

NUMBER OF SEQ ID NOS: 53

; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 23  
; LENGTH: 439  
; TYPE: DNA  
; ORGANISM: Rattus norvegicus  
US-09-956-622A-23

Query Match 71.9%; Score 125.8; DB 12; Length 439;  
Best Local Similarity 96.8%; Pred. No. 2.6e-35;  
Matches 150; Conservative 0; Mismatches 2; Indels 3; Gaps 2;  
  
QY 1 AGTCCAGGAACCTTGAGCTTTGATTTTCAGGAATGCACATCTCTTAGCACTCCCAAAAC 60  
|||||  
Db 285 AGTCCAGGAACCTTGAGCTTTGATTTTCAGGAATGCACATCTCTTAGCACTCCCAAAAC 344  
|||||  
QY 61 AGGAAGGCTCCACACCTCTGCGAGCGCCGCTTTCTTTCAGCATGAGAAACACAAGG 120  
|||||  
Db 345 AGGAAGGCTCCACACCTCTGCGAGCGCCGCTTTCTTTCAGCATGAGAAACACAAGG 404  
|||||  
QY 121 G--ACAGCAGAGTAC-TCTCCTCTGGAGGACTAGT 152  
| |||||  
Db 405 GGACCAGCAGAGTACTTTTCTCTCTGGAGGACTAGT 439  
|||||

RESULT 5  
US-10-027-632-12045/c  
; Sequence 12045, Application US/10027632  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027.632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12045  
; LENGTH: 707  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-12045

Query Match 17.8%; Score 31.2; DB 15; Length 707;  
Best Local Similarity 54.3%; Pred. No. 0.47;  
Matches 63; Conservative 0; Mismatches 53; Indels 0; Gaps 0;  
  
QY 36 CACATCTCTTAGCACTCGCAAAACAGAGGCTCCACACCTCTGGCAGCGCGGCTT 95  
|||||  
Db 686 CTCAGTCCCAGAGTCTCTCAGAGGCTGAGGCTCCCATCTCTGCGAGCAGCGCCTA 627  
|||||  
QY 96 TCTCTCAGCATGAGAAGACAGGACAGCAGAGTACTCTCTCTGAGGACTAG 151  
|| |||||  
Db 626 GCTAACGNGTCACAGAAACAGGCGC?AAGCAAAAGTCCCAATCCAGAGAGGCTG 571  
|||||

RESULT 6  
US-09-796-692-359  
; Sequence 359, Application US/09796692  
; Publication No. US20020198362A1

; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THER  
; TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES  
; FILE REFERENCE: 2077.001200  
; CURRENT APPLICATION NUMBER: US/09/796.692  
; CURRENT FILING DATE: 2001-03-01  
; PRIOR APPLICATION NUMBER: 60/186,126  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: 60/190,479  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: 60/200,545  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: 60/200,303  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: 60/200,779  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: 60/200,999  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: 60/202,084  
; PRIOR FILING DATE: 2000-05-04  
; PRIOR APPLICATION NUMBER: 60/206,201  
; PRIOR FILING DATE: 2000-05-22  
; PRIOR APPLICATION NUMBER: 60/218,950  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: 60/222,903  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: 60/223,416  
; PRIOR FILING DATE: 2000-08-04  
; PRIOR APPLICATION NUMBER: 60/223,378  
; PRIOR FILING DATE: 2000-08-07  
; NUMBER OF SEQ ID NOS: 9597  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 359  
; LENGTH: 463  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(463)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-796-692-359

Query Match 17.6%; Score 30.8; DB 11; Length 463;  
Best Local Similarity 53.3%; Pred. No. 0.57;  
Matches 65; Conservative 0; Mismatches 57; Indels 0; Gaps 0;  
  
QY 1 AGTCCAGGAACCTTGAGCTTTGATTTTCAGGAATGCACATCTCTTAAGCACTCCCAAAAC 60  
|||||  
Db 265 AGTCTTCAACTCTCTGTTTCTGCTAACAGAGGTGTGCAATGAATTACCACTGGTAACC 324  
|||||  
QY 61 AGGAAGGCTCCACACCTCTGCGAGCGCCGCTTTCTTTCAGCATGAGAAACACAAGG 120  
|||||  
Db 325 TTGAATGAGTAAATCAAAATCCAGGCGCATGCTCCACGCCCAACAGTCAAAAGCCCAAGG 384  
121 GA 122  
385 AA 386

RESULT 7  
US-09-796-692-4928  
; Sequence 4928, Application US/09796692  
; Publication No. US20020198362A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THER  
; TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES  
; FILE REFERENCE: 2077.001200

CURRENT APPLICATION NUMBER: US/09/796,692  
CURRENT FILING DATE: 2001-03-01  
PRIOR APPLICATION NUMBER: 60/186,126  
PRIOR FILING DATE: 2000-03-01  
PRIOR APPLICATION NUMBER: 60/190,479  
PRIOR FILING DATE: 2000-03-17  
PRIOR APPLICATION NUMBER: 60/200,545  
PRIOR FILING DATE: 2000-04-27  
PRIOR APPLICATION NUMBER: 60/200,303  
PRIOR FILING DATE: 2000-04-28  
PRIOR APPLICATION NUMBER: 60/200,779  
PRIOR FILING DATE: 2000-04-28  
PRIOR APPLICATION NUMBER: 60/200,999  
PRIOR FILING DATE: 2000-05-01  
PRIOR APPLICATION NUMBER: 60/202,084  
PRIOR FILING DATE: 2000-05-04  
PRIOR APPLICATION NUMBER: 60/206,201  
PRIOR FILING DATE: 2000-05-22  
PRIOR APPLICATION NUMBER: 60/218,950  
PRIOR FILING DATE: 2000-07-14  
PRIOR APPLICATION NUMBER: 60/222,903  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: 60/223,416  
PRIOR FILING DATE: 2000-08-04  
PRIOR APPLICATION NUMBER: 60/223,378  
PRIOR FILING DATE: 2000-08-07  
NUMBER OF SEQ ID NOS: 9597  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 4928  
LENGTH: 463  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (6)  
OTHER INFORMATION: n=A,T,C or G  
NAME/KEY: unsure  
LOCATION: (43)  
OTHER INFORMATION: n=A,T,C or G  
NAME/KEY: unsure  
LOCATION: (45)  
OTHER INFORMATION: n=A,T,C or G  
NAME/KEY: unsure  
LOCATION: (47)  
OTHER INFORMATION: n=A,T,C or G  
NAME/KEY: unsure  
LOCATION: (108)  
OTHER INFORMATION: n=A,T,C or G  
NAME/KEY: unsure  
LOCATION: (110)  
OTHER INFORMATION: n=A,T,C or G  
NAME/KEY: unsure  
LOCATION: (420)  
OTHER INFORMATION: n=A,T,C or G  
NAME/KEY: unsure  
LOCATION: (448)  
OTHER INFORMATION: n=A,T,C or G  
US-09-796-692-4928

Query Match 17.6%; Score 30.8; DB 11; Length 463;  
Best Local Similarity 53.3%; Pred. No. 0.57;  
Matches 65; Conservative 0; Mismatches 57; Indels 0; Gaps 0;  
QY 1 AGTCAGGAGCTTGTAGCTTTGTATTTTTCAGGAATGCACATCTCTTAAGCACTCGCAAAAC 60  
DB 265 AGTCCTCAACTCTGTCTTCTGCTTAACAGGTGTGCAATGAATTAACCACTGGGTAAAC 324  
QY 61 AGGAAGGCTCCACACTCTGGCAGGCCAGGGCTTTCTTTCAGCATGAGAAAGACAAGG 120  
DB 325 TTGAATGAGTTAATCAAAATTCAGGCCATCGTCCAGGCCACCAACAGTCAAAAGGCCAAGG 384  
QY 121 GA 122

DB 385 AA 386.  
US-10-040-862-359  
RESULT 8  
Sequence 359, Application US/10040862  
Publication No. US20030078396A1  
GENERAL INFORMATION:  
APPLICANT: Gaiger, Alexander  
APPLICANT: Algate, Paul A.  
APPLICANT: Mannion, Jane  
APPLICANT: Retter, Marc  
APPLICANT: Corixa Corporation  
TITLE OF INVENTION: Hematological Malignancies  
FILE REFERENCE: 014058-013520US  
CURRENT APPLICATION NUMBER: US/10/040,862  
CURRENT FILING DATE: 2001-11-06  
PRIOR APPLICATION NUMBER: US 60/186,126  
PRIOR FILING DATE: 2000-03-01  
PRIOR APPLICATION NUMBER: US 60/190,479  
PRIOR FILING DATE: 2000-03-17  
PRIOR APPLICATION NUMBER: US 60/200,545  
PRIOR FILING DATE: 2000-04-27  
PRIOR APPLICATION NUMBER: US 60/200,303  
PRIOR FILING DATE: 2000-04-28  
PRIOR APPLICATION NUMBER: US 60/200,779  
PRIOR FILING DATE: 2000-04-28  
PRIOR APPLICATION NUMBER: US 60/200,999  
PRIOR FILING DATE: 2000-05-01  
PRIOR APPLICATION NUMBER: US 60/202,084  
PRIOR FILING DATE: 2000-05-04  
PRIOR APPLICATION NUMBER: US 60/206,201  
PRIOR FILING DATE: 2000-05-22  
PRIOR APPLICATION NUMBER: US 60/218,950  
PRIOR FILING DATE: 2000-07-14  
PRIOR APPLICATION NUMBER: US 60/222,903  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: US 60/223,416  
PRIOR FILING DATE: 2000-08-04  
PRIOR APPLICATION NUMBER: US 60/223,378  
PRIOR FILING DATE: 2000-08-07  
PRIOR APPLICATION NUMBER: US 09/796,692  
PRIOR FILING DATE: 2001-03-01  
NUMBER OF SEQ ID NOS: 10467  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 359  
LENGTH: 463  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc.feature  
LOCATION: (1)-(463)  
OTHER INFORMATION: n = A,T,C or G  
US-10-040-862-359

Query Match 17.6%; Score 30.8; DB 15; Length 463;  
Best Local Similarity 53.3%; Pred. No. 0.57;  
Matches 65; Conservative 0; Mismatches 57; Indels 0; Gaps 0;  
QY 1 AGTCAGGAGCTTGTAGCTTTGTATTTTTCAGGAATGCACATCTCTTAAGCACTCGCAAAAC 60  
DB 265 AGTCCTCAACTCTGTCTTCTGCTTAACAGGTGTGCAATGAATTAACCACTGGGTAAAC 324  
QY 61 AGGAAGGCTCCACACTCTGGCAGGCCAGGGCTTTCTTTCAGCATGAGAAAGACAAGG 120  
DB 325 TTGAATGAGTTAATCAAAATTCAGGCCATCGTCCAGGCCACCAACAGTCAAAAGGCCAAGG 384  
QY 121 GA 122  
DB 385 AA 386





/ PRIOR APPLICATION NUMBER: PCT/US01/00662  
/ PRIOR FILING DATE: 2001-01-30  
/ PRIOR APPLICATION NUMBER: PCT/US01/00661  
/ PRIOR FILING DATE: 2001-01-30  
/ PRIOR APPLICATION NUMBER: PCT/US01/00670  
/ PRIOR FILING DATE: 2001-01-30  
/ PRIOR APPLICATION NUMBER: US 60/234,687  
/ PRIOR FILING DATE: 2000-09-21  
/ PRIOR APPLICATION NUMBER: US 09/608,408  
/ PRIOR FILING DATE: 2000-06-30  
/ PRIOR APPLICATION NUMBER: US 09/774,203  
/ PRIOR FILING DATE: 2001-01-29  
/ NUMBER OF SEQ ID NOS: 49117  
/ SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
/ SEQ ID NO 17772  
/ LENGTH: 1096  
/ TYPE: DNA  
/ ORGANISM: Homo sapiens  
/ FEATURE:  
/ OTHER INFORMATION: MAP TO AF111170.3  
/ OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.7  
/ OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.2  
/ OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2  
/ OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.8  
/ OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.2  
/ OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2  
/ OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 4.9  
/ OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6  
/ OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.6  
/ OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 6  
/ OTHER INFORMATION: NT HIT: AF111170.3, EVALUATE 0.00e+00  
/ OTHER INFORMATION: EST\_HUMAN HIT: BF364900.1, EVALUATE 0.00e+00  
/ OTHER INFORMATION: SWISSPROT HIT: O14682, EVALUATE 1.00e-03  
US-09-864-761-17772

Query Match 17.1%; Score 30; DB 10; Length 1096;  
Best Local Similarity 61.5%; Pred. No. 1.5;  
Matches 48; Conservative 0; Mismatches 30; Indels 0; Gaps 0

QY 43 CTTAAGCACTCGCAAACAGGAAGGCTCCACACTCTGGCAGCGCAGGCCTTTCTCTTC 102  
DB 416 CTCGGGTAGGCAGAGACTGGATCGTGCGGCGAGGCCCTCTCTCTTG 475

QY 103 AGCATGAGAAAGACAAGG 120  
DB 476 GTCAGGGGAAAGTCAGG 493

RESULT 12  
US-10-161-803-23  
Sequence 23, Application US/10161803  
Publication No. US20030092028A1  
GENERAL INFORMATION:  
APPLICANT: Ma, Yuanhong  
APPLICANT: Li, Chih-Jian  
APPLICANT: Chen, Fan  
APPLICANT: Fairman, Jeffery  
TITLE OF INVENTION: METHODS AND REAGENTS FOR DIAGNOSIS AND TREATMENT OF INSULIN RESISTANCE AND RELATED CONDITIONS  
FILE REFERENCE: 42145200300  
CURRENT APPLICATION NUMBER: US/10/161,803  
CURRENT FILING DATE: 2002-06-03  
PRIOR APPLICATION NUMBER: US 60/295,264  
PRIOR FILING DATE: 2001-06-01  
NUMBER OF SEQ ID NOS: 61  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 23  
LENGTH: 1753  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-161-803-23

```
Query Match      16.8%; Score 29.4; DB 15; Length 1753;
Best Local Similarity 58.6%; Pred. No. 2.9;
Matches 51; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

Qy 54 GCAAAACAGGAGGCTCCACACCTCTGGCAGGCCAGGCCCTTCTTTCAGCATGAGAAA 113
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 593 GGACAGCAGCATGTCACACCTGCGCTCCACATCGAGGCCCTCGCGTGGGAGAGCGAGA 652
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 114 GACAAGGACAGCAGAGTACTCTCTC 140
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 653 CCCCTGTGACITCATAGGACCCTACTC 679
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 13
US-10-027-632-205550/c
; Sequence 205550, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 205550
; LENGTH: 624
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-205550

Query Match      16.6%; Score 29; DB 15; Length 624;
Best Local Similarity 52.08; Pred. No. 2.8;
Matches 65; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

Qy 47 AGCACTCGCAAAACAGGAGGCTCCACACCTCTGGCAGGCCAGGCCCTTCTTTCAGCA 106
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 155 AACACTCTAATTTCTCTGAATGCCCTAAACTGTAGCAGACAATGCTCTTAATTATTAACA 96
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 107 TGAGAAACAGGAGGACAGCAGAGTACTCTCTCTGGAGGACTAGTCTAGCCTAGATAA 166
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 95 ACTATTATTACATGGACAGAGATCTTAGGCTTCCAGAAATTAGAAGGGTAACTTAGGACAA 36
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 167 ACACC 171
    |||||
Db 35 AAAAC 31
    |||||

RESULT 14
US-10-027-632-110489
; Sequence 110489, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
```

```
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 110489
; LENGTH: 851
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-110489

Query Match      16.6%; Score 29; DB 15; Length 851;
Best Local Similarity 63.8%; Pred. No. 3.2;
Matches 44; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

Qy 60 CAGGAAGGCTCCACACCTCTGGCAGGCCAGGCCCTTCTTTCAGCATGAGAAAGACAAG 119
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 108 CAGGTGGCTCAAACTCTGACCTAAGTGATCCTCCCTCATCAGCCTCCCAAGTGTG 167
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 120 GGACAGCAG 128
    ||||| |||||
Db 168 GGACTGCAG 176
    ||||| |||||

RESULT 15
US-09-560-863-222/c
; Sequence 222, Application US/09560863
; Patent No. US20020110809A1
; GENERAL INFORMATION:
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Nehls, Michael C.
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. US20020110809A1 Human Polynucleotides and the
; FILE REFERENCE: LEX-0018-USA
; CURRENT APPLICATION NUMBER: US/09/560,863
; CURRENT FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/132,408
; PRIOR FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 1008
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 222
; LENGTH: 431
; TYPE: DNA
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(431)
; OTHER INFORMATION: n = A,T,C or G
; US-09-560-863-222

Query Match      16.5%; Score 28.8; DB 11; Length 431;
Best Local Similarity 62.5%; Pred. No. 2.9;
Matches 45; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

Qy 60 CAGGAAGGCTCCACACCTCTGGCAGGCCAGGCCCTTCTTTCAGCATGAGAAAGACAAG 119
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 312 CAGGGAAGCCACAGCTTCTGGAGACACACGGGCTATGTTCTCTCTAGGAGAAAGACG 253
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 120 GGACAGCAGAGT 131
    ||||| |||||
Db 252 AGTACCAGGCT 241
    ||||| |||||
```

Fri Aug 1 10:44:12 2003

us-10-010-408-1\_copy\_1534\_1708.rnpb

Page 8

Search completed: July 28, 2003, 21:31:58  
Job time : 47.1391 secs

GenCore version 5.1.6  
Copyright (c) 1993.- 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 18:02:17 ; Search time 310.315 Seconds  
(without alignments)  
9133.322 Million cell updates/sec

Title: US-10-010-408-l\_COPY\_1534\_1708

Perfect score: 175

Sequence: 1 AGTCCAGGAAGTGGAGCTTT.....GCCTAGATAAACACCCAAA 175

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estm:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_htc:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_htc:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: gb\_gss:\*

18: em\_gss\_hum:\*

19: em\_gss\_inv:\*

20: em\_gss\_pln:\*

21: em\_gss\_vit:\*

22: em\_gss\_fun:\*

23: em\_gss\_mam:\*

24: em\_gss\_mus:\*

25: em\_gss\_othr:\*

26: em\_gss\_pro:\*

27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|-------|---------------|--------|-------|-------------|
| c 1        | 173.4 | 99.1          | 762    | 14    | BQ195526    |
| 2          | 126.2 | 72.1          | 418    | 9     | A1225477    |
| 3          | 124.6 | 71.2          | 313    | 10    | BB367824    |
| 4          | 123   | 70.3          | 315    | 10    | BB374499    |
| 5          | 119.8 | 68.5          | 307    | 10    | BB220676    |
| 6          | 119.8 | 68.5          | 337    | 10    | BB319151    |

|      |       |      |      |    |          |
|------|-------|------|------|----|----------|
| 7    | 119.8 | 68.5 | 345  | 10 | BB222795 |
| 8    | 113.4 | 64.8 | 281  | 10 | BB309266 |
| 9    | 113.4 | 64.8 | 316  | 10 | BB518921 |
| 10   | 110.2 | 63.0 | 218  | 10 | BB213539 |
| 11   | 110.2 | 63.0 | 289  | 10 | BB186801 |
| 12   | 110.2 | 63.0 | 290  | 10 | BB515213 |
| 13   | 104.8 | 59.9 | 369  | 10 | BB792544 |
| 14   | 100.2 | 57.3 | 324  | 10 | BB558051 |
| 15   | 74.4  | 42.5 | 792  | 12 | BF138093 |
| 16   | 33.6  | 19.2 | 675  | 17 | AZ574808 |
| 17   | 33.6  | 19.2 | 712  | 13 | BI147316 |
| c 18 | 33.4  | 19.1 | 360  | 17 | AQ345221 |
| c 19 | 33.4  | 19.1 | 506  | 10 | BB701454 |
| c 20 | 32.2  | 19.0 | 469  | 9  | AL800293 |
| c 21 | 32.2  | 18.4 | 317  | 12 | BF077992 |
| 22   | 32.2  | 18.4 | 441  | 10 | BB775647 |
| 23   | 32.2  | 18.4 | 589  | 12 | BF712425 |
| c 24 | 32.2  | 18.4 | 620  | 10 | AV751782 |
| 25   | 32    | 18.3 | 677  | 17 | AZ574825 |
| 26   | 32    | 18.3 | 682  | 17 | AZ574814 |
| 27   | 32    | 18.3 | 744  | 17 | CNS04PR1 |
| 28   | 32    | 18.3 | 1089 | 17 | CNS031YH |
| 29   | 31.8  | 18.2 | 790  | 9  | AL820075 |
| c 30 | 31.6  | 18.1 | 799  | 13 | BI757228 |
| 31   | 31.4  | 17.9 | 1135 | 14 | BQ278875 |
| 32   | 31.2  | 17.8 | 412  | 10 | AW694655 |
| 33   | 31.2  | 17.8 | 454  | 9  | AL379783 |
| 34   | 31    | 17.7 | 263  | 9  | AV063522 |
| c 35 | 31    | 17.7 | 327  | 12 | BF549593 |
| 36   | 31    | 17.7 | 373  | 17 | AQ120776 |
| c 37 | 31    | 17.7 | 434  | 13 | BM256378 |
| c 38 | 31    | 17.7 | 506  | 10 | BE485090 |
| 39   | 31    | 17.7 | 639  | 13 | BJ254458 |
| 40   | 31    | 17.7 | 691  | 17 | AQ752684 |
| 41   | 31    | 17.7 | 697  | 17 | AZ127785 |
| 42   | 31    | 17.7 | 740  | 14 | BQ210400 |
| c 43 | 31    | 17.7 | 799  | 17 | AZ185214 |
| 44   | 31    | 17.7 | 929  | 14 | BQ425393 |
| c 45 | 30.8  | 17.6 | 418  | 14 | W97990   |

#### ALIGNMENTS

RESULT 1  
BQ195526/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
COMMENT

BQ195526  
UI-R-CN1-cmq-k-07-0-UI.s1  
BQ195526  
GI:20371077  
EST.  
Norway rat.  
Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
1 (bases 1 to 762)  
Bonaldo, M.F., Lennon, G. and Soares, M.B.  
Normalization and subtraction: two approaches to facilitate gene  
discovery  
Genome Res. 6 (9), 791-806 (1996)  
97044477  
Contact: Soares, MB  
Program for Rat Gene Discovery and Mapping  
University of Iowa  
451 Eckstein Medical  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: msoares@blue.weeg.uiowa.edu  
The sequence contained an oligo-dT track that was present in the  
oligonucleotide that was used to prime the synthesis of first  
strand cDNA and therefore this may represent a bonafide poly A

762 bp  
linear  
EST 30-APR-2002  
Rattus norvegicus cDNA clone  
UI-R-CN1-cmq-k-07-0-UI 3', mRNA sequence.

tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to identify it as a clone from the normalized cervix library cDNA Library Preparation: M.B. Soares Lab  
 Clone distribution: clones will be available through Research Genetics (www.resgen.com)  
 Seq primer: M13 Forward  
 POLRA-yes.

## FEATURES

## Location/Qualifiers

1. .762

/organism="Rattus norvegicus"

/strain="Sprague-Dawley"

/db\_xref="taxon:10116"

/clone="UI-R-CNI-cmq-k-07-0-UI"

/clone\_lib="UI-R-CNI"

/dev\_stage="adult"

/lab\_host="DH10B (Life Technologies)"

/note="Vector: pT7n3D-Pac (Pharmacia) with a modified polylinker; Site\_1: Not I; Site\_2: Eco RI; The UI-R-CNI library is a subtracted library derived from the following pool of seven normalized rat libraries: normalized rat seminal vesicles, normalized rat penis, normalized rat bladder, normalized rat cervix, normalized rat brown adipose, normalized rat fundus, and normalized rat salivary gland. It was constructed according to the procedure described by Bonaldo, Lennon & Soares (Genome Research Genome 6: 791-806, 1996). For construction of the CNI library, plasmid DNA from the pool of normalized libraries was electroporated into competent bacteria for the production of single-stranded circular DNA. This was then used as a tracer in a subtractive hybridization with a driver (PCR amplified inserts from a plasmid DNA template preparation) comprising: a) a pool of about 34,000 clones from the Rat Unigene Set corresponding to plates R-5-AA-NN excluding plates R-5-MM and MN. This pool represented 40% of the final driver population. b) a pool of about 29,000 clones from subtracted libraries CA0 and CAL corresponding to plates R-CA0-AWV through R-CA0-AXS, R-CA0-AZX through R-CA0-BAL, R-CA0-BFE through R-CA0-BHY, R-CA0-BJ5, R-CA0-BKE, R-CA0-BKG-H, R-CA0-BKJ-K, R-CA0-BKP through R-CA0-BKS, R-CA0-BKU-V, R-CA0-BLY through R-CA0-BMA, R-CA0-BMG through R-CA0-BME, R-CA0-BNS, R-CA0-BOB through R-CA0-BOJ, R-CA0-BPA through R-CA0-BPG, R-CA0-BBA through R-CA0-BDA, R-CA0-BHZ through R-CA0-BJF, R-CA0-BJR, R-CA0-BJT through R-CA0-BKB, R-CA0-BKD, R-CA0-BKF, R-CA0-BKI, R-CA0-BKT, R-CA0-BLF, R-CA0-BLH through R-CA0-BLN, R-CA0-BLS, R-CA0-BLU-V, R-CA0-BNR, and R-CA0-BLE. The resulting pool represented 20% of the final driver population. c) a pool of about 15,000 clones from non-normalized libraries CS0s, CT0s, CU0s, CM0s, CX0s and normalized libraries CS0, CT0, CU0, CM0, and CX0 corresponding to plates R-CS0s-CBD through R-CS0s-CBO, R-CT0s-CAM through R-CT0s-CAX, R-CU0s-CBP through R-CU0s-CCA, R-CW0s-CCB through R-CW0s-CCM, R-CX0s-CCN through R-CX0s-CCX, R-CS0-BSD, R-CS0-BTD through R-CS0-BTV, R-CS0-BVM, R-CT0-BTW through R-CT0-BUP, R-CT0-BVN, R-CU0-BUQ through R-CU0-BVL, R-CW0-BVV through R-CW0-BWP, R-CW0-BXN through R-CW0-BXO, R-CX0-BWQ through R-CX0-BXM. The resulting pool represented 5% of the final driver population. d) a pool of about 5,000 clones (1,000 from non-normalized eye library CV0 and 4,000 from normalized eye library CV1) corresponding to plates R-CV0-BRH through R-CV0-BRR, R-CV1-BRS through R-CV1-BSC, R-CV1-BSE through R-CV1-BTC, and R-CV1-BVO through R-CV1-BVU. This pool represented about 5% of the final driver population. e) A pool of about 10,000 clones from subtracted library BS2, BV0 and BV0P (7-9.5 kb cDNA library fraction from rat whole embryo), and BX0 (0.5-7kb cDNA library fraction from rat whole embryo) corresponding to plates R-BS2-BDB through R-BS2-BOF, R-BV0-ANK through R-BV0-ANR, R-BV0P-AOI through R-BV0P-AOB, and R-BX0-AOY through R-BX0-ASH. The resulting pool represented 5% of the final driver population. f) a pool of about 7,000 clones from the seven non-normalized libraries that make up the tracer

including CY0, CZ0, DA0, DB0, DC0, DD0, and DE0 corresponding to plates R-CY0-BXP through R-CY0-BXZ, R-CZ0-BYA through R-CZ0-BYI, R-CZ0-BZB-C, R-DA0-BYJ through R-DA0-BYP, R-DA0-BZD through R-DA0-BZH, R-DB0-BYQ through R-DB0-BZP, R-DC0-BZI through R-DC0-BZQ, R-DC0-CAY through R-DC0-CBA, R-DD0-BZR through R-DD0-CAA, The R-DD0-CBB-C, and R-DE0-CAB through R-DE0-CAL. The resulting pool represented about 10% of the final driver population. g) a pool of about 2,000 clones from the pool of normalized libraries, CN0, that makes up the tracer. The corresponding plates are R-CN0-BKW through R-CN0-BLD, R-CN0-BLG, R-CN0-BLP through R-CN0-BLR, R-CN0-BLT, R-CN0-BLM-X, R-CN0-BMB, and R-CN0-BMF through R-CN0-BML. This pool represented 5% of the final driver population. h) a pool of the 28 most abundant clones in the CN0 pool corresponding to the following addresses: bkx-a-09-0-UI, bkx-b-09-0-UI, bkx-b-11-0-UI, bkx-b-10-0-UI, bkx-d-01-0-UI, bkx-d-06-0-UI, bkx-g-08-0-UI, bkx-h-12-0-UI, bky-a-05-0-UI, bkz-a-06-0-UI, bkz-a-11-0-UI, bkz-c-06-0-UI, bkz-c-09-0-UI, bkz-d-10-0-UI, bla-a-01-0-UI, bla-a-02-0-UI, bla-f-04-0-UI, bla-g-07-0-UI, bla-g-12-0-UI, blb-a-12-0-UI, blb-f-02-0-UI, blc-a-11-0-UI, blc-e-95-0-UI, bld-1-08-0-UI, bld-f-02-0-UI, blg-h-04-0-UI, blr-a-05-0-UI, blt-f-08-0-UI. This pool represented 5% of the final driver population. i) One abundant CN0 clone (corresponding to the address bkz-a-11-0-UI) was digested with Not I and Eco RI and the resulting insert was gel purified. This purified insert was added directly to the driver so that it represented 5% of the final driver population.

TAG\_LIB=UI-R-CNI

TAG\_TISSUE=cervix

TAG\_SEQ=GACCA"

BASE COUNT 177 a 176 c 198 g 211 t  
 ORIGIN

Query Match 99.1%; Score 173.4; DB 14; Length 762;  
 Best Local Similarity 99.4%; Pred. No. 2.6e-48;  
 Matches 174; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 AGTCAGGAACTTGAGCTTTGTATTTCAGGAATGCACATCTCTTAAGCACTCGCAAAAC 60  
 DB 202 AGTCAGGAACTTGAGCTTTGTATTTCAGGAATGCACATCTCTTAAGCACTCGCAAAAC 143  
 QY 61 AGGAAGCTCCACACCTCTGGCAGGCCAGGGCTTTCTTTCAGCATGAGAAGACAAG 120  
 DB 142 AGGAAGCTCCACACCTCTGGCAGGCCAGGGCTTTCTTTCAGCATGAGAAGACAAG 83  
 QY 121 GACAGCAGAGTACTCTCTCTGGAGCACTAGTCTAGCCTAGATAAACACCCAAA 175  
 DB 82 GACAGCAGAGTACTCTCTCTGGAGCACTAGTCTAGCCTAGATAAACACCCAAA 28

## RESULT 2

LOCUS AI225477 418 bp mRNA linear EST 29-OCT-1998  
 DEFINITION ue8b01.y1 Soares\_NMPu Mus musculus cDNA clone IMAGE:1498153 5',  
 mRNA sequence.  
 ACCESSION AI225477  
 VERSION AI225477.1 GI:3808530  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 418)  
 AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dletrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.  
 TITLE The WashU-HMI Mouse EST Project  
 JOURNAL Unpublished (1996)

Contact: Marra M/Mouse EST Project  
 WashU-BHMI Mouse EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: mouseest@wustl.edu  
 This clone is available royalty-free through LLNL ; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 MGI:935757  
 Seq primer: -4ORP from Gibco  
 High quality sequence stop: 403.  
 Location/Qualifiers  
 1..418  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:1498153"  
 /clone\_lib="Soares\_NMPu"  
 /sex="female"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /note="Organ: uterus; Vector: pT7T3D-Pac (Pharmacia) with  
 a modified polylinker; 1st strand cDNA was prepared from  
 pregnant mouse uterus, and was then primed with a Not I -  
 oligo(dT) primer. Double-stranded cDNA was ligated to Eco  
 RI adaptors (Pharmacia), digested with Not I and cloned  
 into the Not I and Eco RI sites of the modified pT7T3  
 vector. Library is normalized. Library was constructed by  
 Bento Soares and M. Fatima Bonaldo."  
 121 a 98 c 100 g 99 t  
 BASE COUNT  
 ORIGIN  
 Query Match 72.1%; Score 126.2; DB 9; Length 418;  
 Best Local Similarity 86.3%; Pred. No. 2.9e-32;  
 Matches 151; Conservative 0; Mismatches 23; Indels 1; Gaps 1;  
 QY 1 AGTCAGGAACCTGTGAGCTTTGATTTTCAGGAATGCACATCTCTTAAAGCACTGCACAAAC 60  
 Ddb | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 225 AGTCAGGAACCTGTGAGCTTTGATTTGTAATAACACATCTCTTAAATGCTCACAAAGC 284  
 QY 61 AGGAAGGCTCCACACCTCGGCAGCCAGCGCTTCTCTTCAGCATGAGAAACAAAGG 120  
 Ddb | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 285 AAG-AGGCTCCACACTTCGGCAGCCAGCGCTTCTCTTCAGCATGAGAGAGACAAAGG 343  
 QY 121 GACAGCAGAGTACTCTCTCTCGAGGACTAGTCTAGCCTAGATAACACCCAAA 175  
 Ddb | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 344 AACAGTAGAGTACCTCTCTCGAGGACTGCGCCGCTCTGGAATAAACCCCAA 398  
 RESULT 3  
 BB367824  
 LOCUS  
 DEFINITION  
 BB367824 RIKEN full-length enriched, 16 days embryo head Mus  
 musculus CDNA clone Cl30039K03 3' similar to AF126063 Mus musculus  
 connective tissue growth factor-like protein precursor (Ctgfl) mRNA  
 , mRNA sequence.  
 BB367824  
 BB367824.1 GI:9079653  
 EST.  
 SOURCE  
 Mouse mouse.  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 313)  
 Konno.H., Aizawa.K., Akahira,S., Akiyama,J., Arakawa,T., Carninci  
 ,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N.,  
 Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M.,  
 Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N.,  
 Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C., Kusakabe,M.,  
 Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y.,  
 Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K., Shibata,K., Shibata  
 ,Y., Shigenoto,Y., Shingawa,A., Shiraki,T., Sogabe,Y., Sugahara,Y.,  
 Suzuki,H., Suzuki,H., Tagawa,A., Takahashi,F., Tomimaga,N., Toyota











GAGAGAGAGATTCTCGATTAAATTAATCCCCCCCCCC 3'. cDNA was cleaved with BamHI and XhoI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I."

BASE COUNT 84 a 82 c 66 g 84 t

ORIGIN

Query Match 64.8%; Score 113.4; DB 10; Length 316; Best Local Similarity 81.7%; Pred. No. 6e-28; Matches 143; Conservative 0; Mismatches 31; Indels 1; Gaps 1;

QY 1 AGTCCAGGAGTTCAGCTTTGTATTTTTCAGGAGTTCACATCTTTAAGCACTCGCAAAAC 60  
 DB 133 AGTCCAGGAGTTCAGCTTTGTATTTTTCAGGAGTTCACATCTTTTAAATGCTTCACAAAGC 192  
 QY 61 AGGAGGCTCCACACTCTGGCAGGCCAGGCGCTTCTTCAGCATGAGAAGACAAGG 120  
 DB 193 CAG-AGGCTCCACACTCTGCCAGGCCAGGCGCTTTTTCAGCATGAGAGACAAGG 251  
 QY 121 GACAGCAGAGTACTCTCTCTGAGGAGTAGTCTAGCCTAGTAATAAACACCCAAA 175  
 DB 252 AACAGTAGAGTACCTCTCTGAGGAGTGGCGCGTCTGGAATAAACACCCAAA 306

RESULT 10

BB213539 218 bp mRNA linear EST 30-JUN-2000

LOCUS BB213539 RIKEN full-length enriched, adult male aorta and vein Mus musculus cDNA clone A530011G24 3' similar to AF126063 Mus musculus connective tissue growth factor-like protein precursor (ctgfl) mRNA, mRNA sequence.

ACCESSION BB213539

VERSION BB213539.1 GI:8878492

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 218)

AUTHORS Konno, H., Aizawa, K., Akahira, S., Akiyama, J., Hara, A., Hayatsu, N., Hirozane, T., Horii, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M., Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N., Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurihara, C., Kusakabe, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y., Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K., Shibata, K., Shibata, Y., Shigenoto, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Tagawa, A., Takahashi, F., Tomimaga, N., Toya, T., Tsunoda, Y., Watahiki, A., Watanabe, S., Yamamura, T., Yamanaka, I., Yano, R., Yasunishi, A., Yokota, T., Yoshida, K., Yoshiki, A., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

TITLE RIKEN Mouse ESTs (Konno, H., et al.)

JOURNAL Unpublished (2000)

COMMENT Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suenhiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-res@gsc.riken.go.jp.  
 URL: http://genome.gsc.riken.go.jp/  
 Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoka, S., Sasaki, N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 Thermotabilization and thermoactivation of thermostable enzymes by trehalose and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)  
 Itoh, M., Kitsunai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J., Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.  
 Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)  
 Carninci, P. and Hayashizaki, Y.  
 High-efficiency full-length cDNA cloning. Methods Enzymol. 303,

19-44 (1999)  
 Please visit our web site (<http://genome.rtc.riken.go.jp>) for further details.

FEATURES

source location/Qualifiers

1..218

/organism="Mus musculus"

/db\_xref="taxon:10090"

/clone="A530011G24"

/clone\_lib="RIKEN full-length enriched, adult male aorta and vein"

/sex="male"

/tissue\_type="aorta and vein"

/dev\_stage="adult"

/lab\_host="DH10B"

/note="Site 1: Sali; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGATTCTCGATTAAATTAATCCCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I."

BASE COUNT 60 a 57 c 45 g 56 t

ORIGIN

Query Match 63.0%; Score 110.2; DB 10; Length 218; Best Local Similarity 30.6%; Pred. No. 6e-27; Mismatches 0; Indels 1; Gaps 1;

Matches 141; Conservative 0;

QY 1 AGTCCAGGAGTTCAGCTTTGTATTTTTCAGGAGTTCACATCTTTAAGCACTCGCAAAAC 60  
 DB 35 AGTCCAGGAGTTCAGCTTTGTATTTTTCAGGAGTTCACATCTTTTCAACCTTCACAAAC 94  
 QY 61 AGGAGGCTCCACACTCTGGCAGGCCAGGCGCTTCTTCAGCATGAGAAGACAAGG 120  
 DB 95 CAG-AGGCTCCACACTCTGGCAGGCCAGGCGCTTTTTCAGCATGAGAGACAAGG 153  
 QY 121 GACAGCAGAGTACTCTCTCTGAGGAGTAGTCTAGCCTAGTAATAAACACCCAAA 175  
 DB 154 AACAGTAGAGTACCTCTCTGAGGAGTGGCGCGTCTGGAATAAACACCCAAA 208

RESULT 11

BB186801

LOCUS BB186801 RIKEN full-length enriched, adult male spinal cord Mus musculus cDNA clone A330037A13 3' similar to AF126063 Mus musculus connective tissue growth factor-like protein precursor (ctgfl) mRNA, mRNA sequence.

ACCESSION BB186801

VERSION BB186801.1 GI:8847372

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 289)

AUTHORS Konno, H., Aizawa, K., Akahira, S., Akiyama, J., Arakawa, T., Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hirozane, T., Horii, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M., Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N., Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurihara, C., Kusakabe, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y., Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K., Shibata, K., Shibata, Y., Shigenoto, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Tagawa, A., Takahashi, F., Tomimaga, N., Toya, T., Tsunoda, Y., Watahiki, A., Watanabe, S., Yamamura, T., Yamanaka, I., Yano, R., Yasunishi, A., Yokota, T., Yoshida, K., Yoshiki, A., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

Unpublished (2000)  
Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel.: 81-45-503-5222  
Fax: 81-45-503-5216

22708

**BASE CO  
ORIGIN**

28

RESOLUTION  
AB515213

ACCESSION

ORGANISM

## AUTHORS

**TITLE**

JOURNAL

COMMENT

Email: [genome-res@asc.riken.go.jp](mailto:genome-res@asc.riken.go.jp)

Itoh, M., Katsunai, T., Akiyama, J.,

Automated filtration-based high-t-

13 44 (1335)  
please visit our web site (<http://>

## FEATURES

sources

```
heart
/tissue type="heart"
```

```
/note="Site_1: Sali; Site
```

Project of Genome Exploration

RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5']  
 GAGAGAGAGCGCGCAACCTCGAGTTTTTTTTTTTTTNN 3', cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5']  
 GAGAGAGATCTCGAGTTAATAATTAATCCCGCCCCCCC 3'. cDNA was cleaved with BamHI and XhoI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FIC I."

BASE COUNT 76 a 83 c 56 g 75 t  
 ORIGIN  
 Query Match 63.0%; Score 110.2; DB 10; Length 290;  
 Best Local Similarity 80.6%; Pred. No. 7.2e-27;  
 Matches 141; Conservative 0; Mismatches 33; Indels 1; Gaps 1;  
 QY 1 AGTCAGGAGTGTGAGCTTTGTTTTCAGGAATGACATCTTTAAGCACTCGCAAAAC 60  
 |||||  
 Db 108 AGTCAGGAGTGTGAGCTTTGTTTTCAGGAATGACATCTTTAAGCACTCGCAAAAC 167  
 |||||  
 QY 61 AGGAGGCTCCACACTCTGCGAGCGAGGCTTTCTTCAGCATGAGAAGACAAAGG 120  
 |||||  
 Db 168 CAG-AGGCTCCACACTCTGCGAGCGAGGCTTTCTTCAGCATGAGAAGACAAAGG 226  
 |||||  
 QY 121 GACAGAGAGTACTCTCTCTGAGGAGTAGTCTAGCCCTAGATAAACACCCAAA 175  
 |||||  
 Db 227 AACAGTAGTAGTACCCTCTCTGAGGAGTGGCCCTCTCTGGAATAAACACCCAAA 281  
 |||||

## RESULT 13

BB792544  
 LOCUS  
 DEFINITION  
 BB792544 RIKEN full-length enriched, kidney CCL-142 RAG CDNA Mus  
 musculus cDNA clone G430141015 3', mRNA sequence.  
 BB792544  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM

EST.  
 house mouse.  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 369)

REFERENCE  
 AUTHORS  
 Akimura, T., Arakawa, T., Carninci, P., Furuno, M., Hanagaki, T., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Imotani, K., Ishii, Y., Ito, M., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Matsuyama, T., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Watahiki, A., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.  
 RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al. 2001)

JOURNAL  
 COMMENT  
 Unpublished (2001)  
 Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-res@gsc.riken.go.jp,  
 URL: http://genome.gsc.riken.go.jp/  
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., and Hayashizaki, Y., M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)  
 waji, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)  
 Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y. and Hayashizaki, Y.  
 Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

e mouse tissues.  
 Location/Qualifiers  
 1..369  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /clone\_lib="G430141015"  
 /clone\_lib="RIKEN full-length enriched, kidney CCL-142 RAG CDNA"  
 /tissue\_type="kidney"  
 /cell\_line="CCL-142 RAG"  
 /note="pooled cell lines ; (cell\_line=CRL-1751 WEHI 164), (cell\_line=CRL-2116 JC), (cell\_line=RCB-0035 WEHI-3), (cell\_line=RCB-0464 Meth-A), (cell\_line=RCB-1283 B16), (cell\_line=RCB-0559 K-1 Fl), (cell\_line=RCB-1702 WEHI 231 melanoma), (cell\_type=B cells, cell\_line=CRL-1702 WEHI 231), (cell\_type=Leydig cells, cell\_line=CRL-2065 MLTC-1), (cell\_type=Nullipotent stem cell, cell\_line=CRL-2070 NE), (tissue\_type=bladder, cell\_line=RCB-0544 MBT-2), (tissue\_type=bone marrow, cell\_type=stroma cell, cell\_line=CRL-2028 SR-4987), (tissue\_type=colon, cell\_line=RCB-0549 C1e-H3), (tissue\_type=kidney, cell\_line=CCL-142 RAG), (tissue\_type=submandibular gland, cell\_line=CRL-1734 SCA-9 clone 15), (strain=BALB/c, cell\_type=B cells, cell\_line=CRL-1669 BCL1 Clone 13.20-3B3), (strain=C3H, tissue\_type=brain, cell\_line=CRL-1443 BC3H1)"

BASE COUNT 102 a 90 c 86 g 91 t  
 ORIGIN

Query Match 59.9%; Score 104.8; DB 10; Length 369;  
 Best Local Similarity 85.8%; Pred. No. 6e-25;  
 Matches 151; Conservative 0; Mismatches 22; Indels 3; Gaps 3;

QY 1 AGTCAGGAGTGTGAGCTTTGTTTTCAGGAATGACATCTTTAAGCACTCGCAAAAC 60  
 |||||  
 Db 187 AGTCAGGAGTGTGAGCTTTGTTTTCAGGAATGACATCTTTAAGCACTCGCAAAAC 246  
 |||||  
 QY 61 AGGAGGCTCCACACTCTGCGAGCGAGGCTTTCTTCAGCATGAGAAGACAAAGG 120  
 |||||  
 Db 247 CAG-AGGCTCCACACTCTGCGAGCGAGGCTTTCTTCAGCATGAGAAGACAAAGG 304  
 |||||  
 QY 121 GACAGCAGAGTACTCTCTCTGAGGAG-CTAGCTCTAGCCTAGATAAACACCCAAA 175  
 |||||  
 Db 305 AACACGACAGTAGTACCTCTCTGAGGAGCTGCCCGGTCTGGAATAAACACCCAAA 360  
 |||||

RESULT 14  
 BB558051  
 LOCUS  
 DEFINITION  
 BB558051 RIKEN full-length enriched, 2 days pregnant adult female ovary Mus musculus cDNA clone E330032A10 3', similar to Afl26063 Mus (Ctgrfl) mRNA, mRNA sequence.

ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 EST.  
 house mouse.  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 324)  
 Konno, H., Aizawa, K., Akahira, S., Akiyama, J., Arakawa, T., Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N.,

| TITLE  | JOURNAL                         | COMMENT                           |
|--|---------------------------------|-----------------------------------|
| 1. The Role of the Teacher in the Classroom                  | Journal of Educational Research | 1980, Vol. 83, No. 1, pp. 1-10    |
| 2. The Impact of Technology on Education                     | Journal of Educational Research | 1980, Vol. 83, No. 2, pp. 11-20   |
| 3. The Importance of Parental Involvement                    | Journal of Educational Research | 1980, Vol. 83, No. 3, pp. 21-30   |
| 4. The Effect of Teacher Expectations on Student Achievement | Journal of Educational Research | 1980, Vol. 83, No. 4, pp. 31-40   |
| 5. The Role of the School in the Community                   | Journal of Educational Research | 1980, Vol. 83, No. 5, pp. 41-50   |
| 6. The Impact of Teacher Education on Student Achievement    | Journal of Educational Research | 1980, Vol. 83, No. 6, pp. 51-60   |
| 7. The Importance of Teacher Evaluation                      | Journal of Educational Research | 1980, Vol. 83, No. 7, pp. 61-70   |
| 8. The Effect of Teacher Salary on Student Achievement       | Journal of Educational Research | 1980, Vol. 83, No. 8, pp. 71-80   |
| 9. The Role of the School in the Community                   | Journal of Educational Research | 1980, Vol. 83, No. 9, pp. 81-90   |
| 10. The Impact of Teacher Education on Student Achievement   | Journal of Educational Research | 1980, Vol. 83, No. 10, pp. 91-100 |

**FEATURES**

BASE COUNT:  
ORIGIN

|    |                       |   |                    |            |             |
|----|-----------------------|---|--------------------|------------|-------------|
|    | Query Match           | 57.3%   | Score 100.2;       | DB 10;     | Length 324; |
|    | Best Local Similarity | 77.5%;  | Pred. No. 2.1e-33; |            |             |
|    | Matches 134;          | Conservative  | 0; Mismatches      | 38; Indels | 1; Gaps 1;  |
| OY | 3                     | TCCAGGAACATTGACTTTGTATTTTCAGGAATGCACATCTCTTAAAGCACTCGCAAACAG  | 62                 |            |             |
| Dd | 145                   | TTCAGCATCTCCACCATTAGTAGTTCTCTTAATAACATATCTCTTTAATCTTCACAACCAA | 204                |            |             |

Search completed: July 29, 2003, 02:51:34  
Job time : 316.515 secs





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 15:59:01 ; Search time 1764.82 seconds  
(without alignments)  
10471.464 Million cell updates/sec

Title: US-10-010-408-1-copy\_1\_635

Perfect score: 635

Sequence: 1 GAGCGTCTGATCTCCAGAG.....GTGATGACGGTGGCTTCACC 635

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.:

- 1: gb\_ba.:
- 2: gb\_htg.:
- 3: gb\_in.:
- 4: gb\_ov.:
- 5: gb\_ov.:
- 6: gb\_pat.:
- 7: gb\_ph.:
- 8: gb\_pl.:
- 9: gb\_pr.:
- 10: gb\_ro.:
- 11: gb\_sts.:
- 12: gb\_sy.:
- 13: gb\_un.:
- 14: gb\_vt.:
- 15: em\_ba.:
- 16: em\_fun.:
- 17: em\_hum.:
- 18: em\_in.:
- 19: em\_mu.:
- 20: em\_om.:
- 21: em\_or.:
- 22: em\_ov.:
- 23: em\_pat.:
- 24: em\_ph.:
- 25: em\_pi.:
- 26: em\_ro.:
- 27: em\_sts.:
- 28: em\_un.:
- 29: em\_vt.:
- 30: em\_htg\_hum.:
- 31: em\_htg\_inv.:
- 32: em\_htg\_other.:
- 33: em\_htg\_mus.:
- 34: em\_htg\_pln.:
- 35: em\_htg\_rod.:
- 36: em\_htg\_mam.:
- 37: em\_htg\_vrt.:
- 38: em\_sy.:
- 39: em\_htgo\_hum.:
- 40: em\_htgo\_mus.:
- 41: em\_htgo\_other.:

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description          |
|------------|-------|-------------|--------|-------|----------------------|
| 1          | 581.4 | 91.6        | 1741   | 10    | AF259981 Rattus no   |
| 2          | 521.4 | 82.1        | 1734   | 6     | AR210324 Sequence    |
| 3          | 521.4 | 82.1        | 1734   | 6     | AR210325 Sequence    |
| 4          | 521.4 | 82.1        | 1734   | 10    | AF100778 Mus muscu   |
| 5          | 513.6 | 80.9        | 1739   | 10    | AF126063 Mus muscu   |
| 6          | 273   | 43.0        | 1266   | 6     | AX076919 Sequence    |
| 7          | 273   | 43.0        | 1266   | 6     | AX464186 Sequence    |
| 8          | 273   | 43.0        | 1283   | 9     | AF083500 Homo sapi   |
| 9          | 273   | 43.0        | 1293   | 6     | AR210322 Sequence    |
| 10         | 273   | 43.0        | 1293   | 6     | AR210323 Sequence    |
| 11         | 273   | 43.0        | 1309   | 9     | AF074604 Homo sapi   |
| 12         | 273   | 43.0        | 1427   | 9     | AF100780 Homo sapi   |
| 13         | 273   | 43.0        | 1450   | 9     | BC017782 Homo sapi   |
| 14         | 271.4 | 42.7        | 841    | 6     | AR210338 Sequence    |
| 15         | 261.4 | 41.2        | 738    | 6     | AR210337 Sequence    |
| 16         | 221.6 | 34.9        | 137964 | 2     | AC126895 Rattus no   |
| 17         | 221.6 | 34.9        | 226303 | 2     | AC095418 Rattus no   |
| 18         | 215.4 | 33.9        | 137964 | 2     | AC126895 Rattus no   |
| 19         | 199.2 | 31.4        | 61072  | 10    | AL731698 Mouse DNA   |
| 20         | 199.2 | 31.4        | 216757 | 2     | AL669906 Mus muscu   |
| 21         | 164   | 25.8        | 107260 | 9     | AL139352 Human DNA   |
| 22         | 94.2  | 14.8        | 1987   | 10    | AF218568 Rattus no   |
| 23         | 92.8  | 14.6        | 1805   | 5     | CHKCEF               |
| 24         | 91    | 14.3        | 1871   | 10    | AB015877 Rattus no   |
| 25         | 85.6  | 13.5        | 1146   | 6     | E13814 cDNA encodi   |
| 26         | 85.6  | 13.5        | 1887   | 9     | AF031385 Homo sapi   |
| 27         | 85.6  | 13.5        | 1935   | 9     | AF003594 Homo sapi   |
| 28         | 85.6  | 13.5        | 1985   | 9     | BC001271 Homo sapi   |
| 29         | 85.6  | 13.5        | 2016   | 6     | AX336386 Sequence    |
| 30         | 85.6  | 13.5        | 2016   | 6     | AX354335 Sequence    |
| 31         | 85.6  | 13.5        | 2016   | 9     | HSU62015 Homo sapien |
| 32         | 85.6  | 13.5        | 2021   | 9     | HSIG1PRT             |
| 33         | 85.6  | 13.5        | 2025   | 9     | HSCYR61M             |
| 34         | 85.6  | 13.5        | 2041   | 9     | BC016952 Homo sapi   |
| 35         | 85.6  | 13.5        | 2046   | 9     | BC009199 Homo sapi   |
| 36         | 85.6  | 13.5        | 2052   | 9     | HSCYR61              |
| 37         | 84.8  | 13.4        | 1888   | 10    | AF228049 Rattus no   |
| 38         | 84.2  | 13.3        | 1743   | 5     | AF320592 Xenopus l   |
| 39         | 84    | 13.2        | 1418   | 6     | AX206704 Sequence    |
| 40         | 83.8  | 13.2        | 1480   | 6     | AX206702 Sequence    |
| 41         | 83.8  | 13.2        | 2018   | 10    | MUSCYR61A            |
| 42         | 82.4  | 13.0        | 1585   | 5     | NVI271167 Notophtha  |
| 43         | 82.4  | 13.0        | 1973   | 9     | HSNOVH               |
| 44         | 82.4  | 13.0        | 2270   | 6     | AX035239 Sequence    |
| 45         | 82.4  | 13.0        | 2270   | 6     | AX342093 Sequence    |

ALIGNMENTS

RESULT 1

AF259981

LOCUS

DEFINITION

Rattus norvegicus CCN family protein COP-1 (Cop-1), mRNA, complete cds.

ACCSSION

AF259981

VERSION

AF259981.1

KEYWORDS

SOURCE

ORGANISM

Rattus norvegicus.

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

REFERENCE

1 (bases 1 to 1741)

AUTHORS

Zhang, R., Averboukh, L., Zhu, W., Zhang, H., Jo, H., Dempsey, P.J.,

AF259981 1741 bp mRNA linear ROD 09-MAY-2000  
Rattus norvegicus CCN family protein COP-1 (Cop-1), mRNA, complete cds.

AF259981 GI:7739780

Rattus norvegicus.

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 1741)

Zhang, R., Averboukh, L., Zhu, W., Zhang, H., Jo, H., Dempsey, P.J.,

Coffey, R.J., Pardee, A.B. and Liang, P.  
Identification of rCop-1, a new member of the CCN protein family,  
as a negative regulator for cell transformation  
Mol. Cell. Biol. 18 (10), 6131-6141 (1998)  
98414629  
9742130  
PUBMED  
2 (bases 1 to 1741)  
REFERENCE  
Liang, P.  
Direct Submission  
Submitted (24-APR-2000) Cell Biology, Vanderbilt-Ingram Cancer  
Center, 649 MRB II, Nashville, TN 37232, USA  
FEATURES  
Source  
1..1741  
Location/Qualifiers  
/organism="Rattus norvegicus"  
/db\_xref="taxon:10116"  
1..1741  
/gene="Cop-1"  
262..1014  
/gene="Cop-1"  
/note="secreted protein"  
/codon\_start=1  
/product="CCN family protein COP-1"  
/protein\_id="AAF6901.1"  
/db\_xref="GI:7739781"  
/translation="MRGSLIRLLATSLFLLSMYCAQLRTPCTPWTPOCPQGV  
LVLDGCGCKVCAKRLTSCHEHLVCEPSQGLVCPQAGPGHGVACVLLDDEGCEV  
NGRRYLDGETFKPNCRLVLCRDGGFTCLPLCSEDTLPWDCPRKRIQVPGKCPE  
WYCDGVTPAIQSAOAGHOLSALVTPASADAPWENWSTANGPCSTTCGLGIATRVSN  
QNRFLQLEIQRRLCLPRCLLAARSHSSWNSAF"  
BASE COUNT 386 a 491 c 480 g 384 t  
ORIGIN  
Query Match 91.6%; Score 581.4; DB 10; Length 1741;  
Best Local Similarity 97.8%; Pred. No. 1.4e-129;  
Matches 621; Conservative 0; Mismatches 11; Indels 3; Gaps 3;  
QY 1 GACGCTTCGTATCTCCAGAGGACCTGGGGTGGGACAGGGGCTTGGCAAGGCTCGACCC 60  
DB 17 GACGCTTCGTATCTCCAGAGGACCTGGGGTGGGACAGGGGCTTGGCAAGGCTCGAC 74  
QY 61 GCTGGGAGTGGCTGGGAATGAGGCTTTATTACTGGGAAGTGGAGGCTTAAGAGGCTC 120  
DB 75 GCTGGGAGTGGCTGGGAATGAGGCTTTATTACTGGGAAGTGGAGGCTTAAGAGGCTC 134  
QY 121 CTGTCAGCTTGTCTAAAGTCTTAGCACTTGTGGTGGCTTGGGCTTCACACTGTCTAGA 180  
DB 135 CTGTCAGCTTGTCTAAAGTCTTAGCACTTGTGGTGGCTTGGGCTTCACACTGTCTAGA 194  
QY 181 CACCTTCGTGGGCTCCAGGGCTCAGCTTCAAGTGGGCTTGGGCTTCACACTGTCTAGA 240  
DB 195 CACCTTCGTGGGCTCCAGGGCTCAGCTTCAAGTGGGCTTGGGCTTCACACTGTCTAGA 253  
QY 241 ACAGTGACATGAGGGGACCCACCTGATCTTCTGGCCACTTCTCTCTCTCTCTCTCTC 300  
DB 254 ACAGTGACATGAGGGGACCCACCTGATCCGCTTCTGGCCACTTCTCTCTCTCTCTCTC 313  
QY 301 TCTCAATGTGTGTGCCAGCTGTGCCGACACCTGTACCTGTCTTGGACACACCCCC 360  
DB 314 TCTCAATGTGTGTGCCAGCTGTGCCGACACCTGTACCTGTCTTGGACACACCCCC 373  
QY 361 AGTGCCACAGGGGTACCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 420  
DB 374 AGTGCCACAGGGGTACCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 433  
QY 421 GGAGGCTGGGGAGTCTTGGACACCTGCATGTCTGGACCCACCCAGCCGAGGCTGTGT 480  
DB 434 GGAGGCTGAGGAGTCTTGGACACCTGCATGTCTGGACCCACCCAGCCGAGGCTGTGT 493  
QY 481 GTACGCTGGGGAGGCTTGGGGCCATGGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 540  
DB 494 GTACGCTGGGGAGGCTTGGGGCCATGGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 553  
QY 541 GTAGCTGTGAGGTGAATGGCCGACAGGTACCTGGATGGAGAGACCTTTAAACCCAAATGCA 600

554 GTGACTGTGAGTGAATGGCCGAGTACCTGGATGGAGAGACCTTTAAACCCAAATGCA 613  
QY 601 GGGTCTGTGGCGCTGTGATGACGGTGGCTTCACC 635  
DB 614 GGGTCTGTGGCGCTGTGATGACGGTGGCTTCACC 648  
RESULT 2  
AR210324 1734 bp DNA linear PAT 20-JUN-2002  
LOCUS Sequence 17 from patent US 6387657.  
DEFINITION AR210324  
ACCESSION AR210324  
VERSION AR210324.1 GI:21512525  
KEYWORDS  
SOURCE Unknown...  
ORGANISM Unknown.  
Unclassified.  
REFERENCE 1 (bases 1 to 1734)  
AUTHORS Botstein, D.A., Cohen, R.L., Goddard, A.D., Gurney, A.L., Hillan, K.J.,  
Lawrence, D.A., Levine, A.J., Pennica, D., Roy, M. Ann. and Wood, W.I.  
TITLE WISP polypeptides and nucleic acids encoding same  
JOURNAL Patent: US 6387657-A 17-14-MAY-2002;  
FEATURES Location/Qualifiers  
1..1734  
/organism="unknown"  
BASE COUNT 355 a 491 c 495 g 393 t  
ORIGIN  
Query Match 82.1%; Score 521.4; DB 6; Length 1734;  
Best Local Similarity 92.0%; Pred. No. 3.8e-115;  
Matches 586; Conservative 0; Mismatches 41; Indels 10; Gaps 3;  
QY 3 CGCTTCTGATCTCCAGAGGACCTGGGGTGGGACAGGGGCTTGGCAAGGCTGCAGCGC 62  
DB 13 CGCTTCTGATCTCCAGAGGACCCCGGCTGGGACAGGGGCTTGGGAGGCTGCAGCTGC 72  
QY 63 TG-GCAGCTGGGTTGGAATGAGGCTTTATTACTTGGGAACCTAGAGAGCTAAGAGCTCC 121  
DB 73 TGTGGCAGTAGGTTGGGATGAGGCTTTCTTGTGGGACTGAGAGCTGAGAGCTGCC 132  
QY 122 TGTGAG-CTTGTGCTTAAAGTCTTAGCACTTGTGGTGGCTTGGGCTTCACACTGTGCA 178  
DB 133 TGTGAGGCTGCTGCTTAAACTTGTGGCACTTGGGCTTGGGCTTCACACTGTGCA 192  
QY 179 GACACTTCTGCTGGGCTCCAGCGCTCAGCTTCAAGTGTGCAAGCTGGCTCCACAGG 238  
DB 193 GACACTTCTGCTGGGCTTCTCGGCC-----TCAGGTTTGAAGCTGGCTCCACAGG 246  
QY 239 ACAGGCTGACATGAGGGGACGCCACTGATCCATCTTCTGGCCACTTCTCTCTCTCTCT 298  
DB 247 ACAGGCTGACATGAGGGGACGCCACTGATCCATCTTCTGGCCACTTCTCTCTCTCTCAT 306  
QY 299 TCTCTCAATGAGTGTGGCCAGCTGTGGCGGACACCCCTGTACCTGTCTCTTGGACACACC 358  
DB 307 TCTCTCAATGAGTGTATCCAGCTGTGCCAGCACCCCTGTGCTCTCTCTTGGACACACC 366  
QY 359 CCAGTGCCACAGGGGTACCCCTGCTGCTGAGTGGCTGTGGCTCTCTCTCTCTCTCTCTCT 418  
DB 367 CCAGTGCCACAGGGGTACCCCTGCTGCTGAGTGGCTGTGGCTCTCTCTCTCTCTCTCTCT 426  
QY 419 ACAGGCTGAGGAGTCTTGGGACACCTGCTGCTGCAAGCTTGGCAAGGCTTGGT 478  
DB 427 ACAGGCTGAGGAGTCTTGGGACACCTGCTGCTGCAAGCTTGGCAAGGCTTGGT 486  
QY 479 TGTGAGCTTGGGAGGCTTGGGCGCCATGGGCTGTGTCTCTCTCTCTCTCTCTCTCTCTCT 538  
DB 487 TGTGAGCTTGGGAGGCTTGGGCGCCATGGGCTGTGTCTCTCTCTCTCTCTCTCTCTCTCT 546  
QY 539 CGGTAGCTGTGAGGTGAATGCCGAGGTACCTGGATGGAGAGACCTTTAAACCCAAATG 598  
DB 547 CGGAGCTGTGAGGTGAATGCCGAGGTACCTGGATGGGAGAGACCTTTAAACCCAAATG 606

7/5 GACACCTTCGTTGGTGGCCCTCCACGGCCCTCACCCTTCAGGTTTGAAGCTGGCTCCACAAGGG 238

Y  
1/3 GACACCTTCGGTGGGCGCCCTCCACGGCCCTCAGGCTTCAGGCTTGAAGCTGGCTCC

Db 193 GACACCTCTTGGTGGCCCTCTCGGCC-----TCAGGTTTGAAGTGCTGCCACCAAGG 246  
QY 239 ACACGGTGACATGAGGGGAGCCCACTGATCCATCTTGGCCACTTCTTCTCTGCTGCT 298  
Db 247 ACACGGTGACATGAGGGGAGCCCACTGATCCATCTTGGCCACTTCTTCTCTGCTGCT 306  
QY 299 TCTCTCAATGTTGTGGCCAGCTGTGGCCGACACCCCTGTACTCTTCTTGGACACCAACC 358  
Db 307 TCTCTCAATGTTGTATCCAGCTGTGGCCAGACCCCTGTGCTCTTGGACACCAACC 366  
QY 359 CCAGTGGCCACAGGGGTACCCCTGTGTGATGGCTGTGGCTGTCTTAAAGTGTGTGC 418  
Db 367 CCAGTGGCCACAGGGGTACCCCTGTGTGATGGCTGTGGCTGTCTTAAAGTGTGTGC 426  
QY 419 ACGGAGGCTGGGGAGTCTCTGGACACCTGCTGCGACCCAGCCAGCCAGGGGCTGTGT 478  
Db 427 ACGGAGGCTGGGGAGTCTCTGGACACCTGCTGCGACCCAGCCAGGGGCTGTGT 486  
QY 479 TTGTGAGCTGGGGAGCCCTGGGGGAGCCATGGGCTGTGTCTTCTTGGATGAGATGA 538  
Db 487 TTGTGAGCTGGGGAGCCCTGGGGGAGCCATGGGCTGTGTCTTCTTGGATGAGATGA 546  
QY 539 CGGTAGCTGTAGGTGAATGGCCGAGGTACCTGATGGAGAGACCTTTAAACCCAAATG 598  
Db 547 CGGGAGCTGTAGGTGAATGGCCGAGGTACCTGATGGAGAGACCTTTAAACCCAAATG 606  
QY 599 CAGGTCCTGTGCCGCTGTGATGACGGTGGCTTACC 635  
Db 607 CAGGTCCTGTGATGACGGTGGCTTACC 643

RESULT 5  
AF126063  
LOCUS  
DEFINITION Mus musculus connective tissue growth factor-like protein precursor  
(Ctfl) mRNA, complete cds.  
ACCESSION AF126063  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

AF126063.1 GI:4337059

Mus musculus  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
1 (bases 1 to 1739)  
Kumar, S., Hand, A.T., Connor, J.R., Dodds, R.A., Ryan, P.J.,  
Trill, J.J., Fisher, S.M., Nuttall, M.E., Lipshutz, D.B., Zou, C.,  
Hwang, S.M., Votta, B.J., James, I.E., Riegan, D.J., Gowen, M. and  
Lee, J.C.  
Identification and cloning of a connective tissue growth  
factor-like cDNA from human osteoblasts encoding a novel regulator  
of osteoblast functions  
J. Biol. Chem. 274 (24), 17123-17131 (1999)  
99287915  
10358067  
2 (bases 1 to 1739)  
Kumar, S. and Zou, C.  
Direct Submission  
Submitted (04-FEB-1999) Bone & Cartilage Biology, UW 2109,  
SmithKline Beecham, 709 Swedeland Rd., King of Prussia, PA 19406,  
USA

TITLE  
factor-like cDNA from human osteoblasts encoding a novel regulator  
of osteoblast functions  
J. Biol. Chem. 274 (24), 17123-17131 (1999)  
99287915  
10358067

JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
Source  
Location/Qualifiers  
1..1739  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/tissue\_type="lung"  
1..1739  
/gene="Ctfl"  
242..997  
/gene="Ctfl"  
/note="similar to the Mus musculus Wisp-2 protein encoded  
by the sequence presented in GenBank Accession Number  
AF100778; putative growth factor; CTGF-L; contains IGF  
binding (IGFBP), Von Willebrand Factor type C (VWC) repeat

gene

CDS

RESULT 6  
AX076919  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS

AX076919  
Sequence 31 from Patent WO0105836.  
AX076919  
AX076919.1 GI:13121575

1266 bp  
DNA  
linear

PAT 22-FEB-2001

and thrombospondin type I (TSP1) domains; member of the  
CCN (CTGF/Cyrl61/Nov) family; lacks the fourth  
carboxy-terminal (CT) domain present in other members of  
the CCN family"  
/codon\_start=1  
/product="connective tissue growth factor-like protein  
precursor"  
/protein\_id="AAD18058.1"  
/db\_xref="GI:4337060"  
/translation="MRGNPLHLHLAIFLILSWYSLCPAPCAPWTPPOCPGVP  
LYLDCGGCCVRCARLGGESCDLHVCDFPSOGLVCPGAPSGRGAFLFEEDDGSCEV  
NGRRYLDGETFAPNCRVLCRDGDTCLPLCSEDEVRLPSPWDCPRRRIQVPRCCPE  
WYCDQVQPAIQSPSSAQHQLSALVTPASADGPCPNMSTAMGPCSTTCGLGIATRV  
NONRFOLEIORRLCLSRPCLASRSHGWSNFA"  
BASE COUNT 375 a 480 c 489 g 395 t

Query Match 80.9%; Score 513.6; DB 10; Length 1739;  
Best Local Similarity 91.5%; Pred. No. 2.9e-113;  
Matches 580; Conservative 0; Mismatches 44; Indels 10; Gaps 3;

QY 6 TTCTGATCTCCAGAGACCCCTGGGTGGGACAGGGCCCTTGGCAAGGCTGCAGCCGCTG- 64  
Db 1 TCCTGATCTCCAGAGACCCCGGCTGGGACAGGGCCCTTGGCAGGCTGCAGCTGCTGT 60  
QY 65 GGCAGTGGCTTGGAAATGGAGGTCTTTATTACTTGGAACTGAGGAGCTAAGAGGCTCCTGT 124  
Db 61 GGCAGTGGCTTGGGATGGAGGTCTTTCTTGTCTGGAACTGAGGAGCTGAGAAGCTCCTGT 120  
QY 125 CAG---CTTGTCTTAAAGTCTTAGACACTGTGGTGGCTTGGCTTGCACACACTGTCAGAC 181  
Db 121 CAGGCTCTCTTAAACTCTTGGCACTTGGCGTGGCTTGGCTTGCACACACTGTCAGAC 180  
QY 182 ACCTTCTGTGGTGGCTTCCACAGCCCTCACCTTCAGTGTGGAAGCTGGCTCCACAAAGGACA 241  
Db 181 ACCTTCTGTGGTGGCTTCCCTCGGCC-----TCAGGTTGGAAGCTGGCTCCACAAAGGACA 234  
QY 242 CGGTGACATGAGGGGACGCCACTGATCCATCTTCTGGCCACTTCTTCTCTGCTGCTTCT 301  
Db 235 CGGTGACATGAGGGGACGCCACTGATCCATCTTCTGGCCATTTCTTCTCTGCTGCTTCT 294  
QY 302 CTCATGTTGTGTGCCAGCTGTGGGACACCCCTGTACTCTCTTGGACACCAACCCCA 361  
Db 295 CTCATGTTGTATTCCAGCTGTGCCAGACCCCTGTGCTCTTGGACACCAACCCCA 354  
QY 362 GTGCCACAGGGGGTACCCCTGGTGTGATGGCTGTGGCTGTCTGTAAGTGTGTGCAGG 421  
Db 355 GTGCCACAGGGGGTACCCCTGGTGTGATGGCTGTGGCTGTCTGCTGCTGCTGCTGCTG 414  
QY 422 GAGGCTGGGGAGTCTTGGGACCACTGATCTGTGGACCCAGCCAGCCAGGGGCTGTGTTG 481  
Db 415 GAGGCTGGGGAGTCTTGGGACCACTGATCTGTGGACCCAGCCAGGGGCTGTGTTG 474  
QY 482 TCAGCTGGGGAGGCTTGGGCGCCATGGGCTGTGTCTCTTGGATGAGATGACCGG 541  
Db 475 TCAGCTGGGGAGGCTTGGGCGCCAGTGTGTGCTGTGTGCTTCTTGAAGAGATGACCGG 534  
QY 542 TAGCTGTGAGGTGAATGGCCGAGGTACCTGATGGAGAGAGACTTTAAACCAATTCGAG 601  
Db 535 GAGCTGTGAGGTGAACGGCCGAGGTACCTGATGGAGAGAGACTTTAAACCAATTCGAG 594  
QY 602 GGTCCTGTGCCCTGTGATGACGGTGGCTTACC 635  
Db 595 GGTTTGTGCCCTGTGATGACGGTGGTTCACC 628

|                           |   |  |          |                        |
|---------------------------|---|--|----------|------------------------|
| BASE COUNT                | 216 a   | 418 c  | 390 g    | 242 t                  |
| ORIGIN                    |   |  |          |                        |
| Query Match               | 43.0%   | Score 273:   | DB 6:    | Length 1266;           |
| Best Local Similarity     | 80.9%   | Pred. No.  | 2.5e-55; |                        |
| Matches 318; Conservative | 0;  | Mismatches   | 75;      | Indels 0; Gaps 0;      |
| QY                        | 243   | GGTGACATGAGGGCAGCCACCTGCATCATCTTCTGGCCCACTTCCTTTCCCTCTCGCTTCTC | 302      |                        |
| Db                        | 4   | GGGACATGAGAGCCACACCAGAACCCACCTCTCGGCCCTTCTCCCTCTCTCGCTCCTC     | 63       |                        |
| QY                        | 303   | TCAATGGTGTGCCAGCTGTGCCGAGACACCTGTACCTGTACCTGTGGACACACCCCAG     | 362      |                        |
| Db                        | 64  | TCAAAGGTGGGTACCCAGCTGTGCCGAGACACCATGTACCTGCCCTGGCCACCTGCCCGA   | 123      |                        |
| QY                        | 363   | TGCCCCACAGGGGTACCCCTGGTGTCTGGATGGCTGTGGCTGCTGTAAAGTGTGTGCACGG  | 422      |                        |
| Db                        | 124   | TGCCCGCTGGGATGACCCCTGGTGTGGATGGCTGTGGCTGCTGCCGGGTATGTGCACGG    | 18       |                        |
| QY                        | 423   | AGGCTGGGGAGTCTTCGCGACCACTGCATGTCTGCGCACCCAGCCAGGCGCTGGTTTGT    | 482      |                        |
| Db                        | 184   | CGGCTGGGGAGCCCTGCGACAACCTCCACGTCTGCGACGCCAGGCGCTGTCTGC         | 243      |                        |
| QY                        | 483   | CAGCCTGGGGCAGGCCCTTGGCGCCATGGGGCTGTGTCTTGTGGATGAGGATGACGGT     | 542      |                        |
| Db                        | 244   | CAGCCCGGGCAGGACCCCGTGGCGGGGGCCCTGTGCCTCTTTGGCAGAGGACGACAGC     | 303      |                        |
| QY                        | 543   | AGCTGTGAGTGAATGGCCGAGCTACCTGGATGGAGAGACCTTTAAACCAATTGCAGG      | 602      |                        |
| Db                        | 304   | AGCTGTGAGTGAACGCGCCCTGTATCGGGAAGGGAGAGACCTTCCAGCCCCACTGCAGC    | 363      |                        |
| QY                        | 603   | GTCCTGTGCCGTGTGATGACGGTGGCTTCCACC                              | 635      |                        |
| Db                        | 364   | ATCCCTGCCGCTGCGAGGACGGCGCTTCCACC                               | 396      |                        |
| RESULT 8                  |   |  |          |                        |
| AF083500                  |   |  |          |                        |
| LOCUS                     |   |  |          |                        |
| DEFINITION                | Homo sapiens connective tissue growth factor-like protein precursor, mRNA, complete cds.  | 1283 bp  | mRNA     | linear PRI 04-NOV-1998 |
| ACCESSION                 | AF083500  |  |          |                        |
| KEYWORDS                  | AF083500.1 GI:3462835   |  |          |                        |
| SOURCE                    | Homo sapiens.   |  |          |                        |
| ORGANISM                  | Homo sapiens  |  |          |                        |
| REFERENCE                 | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.   |  |          |                        |
| AUTHORS                   | Kumar,S., Hand,A.T., Connor,J.C., Dodds,R.A., Ryan,P.J., Trill,J.G., Fisher,S.M., Slemmon,J.R., Lipshutz,D.B., Bartholomew,I.V., James,I.E., Rieman,D.J., Gowen,M. and Lee,J.C. |  |          |                        |
| TITLE                     | Identification and cloning of CTGF-L from human osteoblasts, a novel cysteine rich protein containing an IGF binding domain   |  |          |                        |
| JOURNAL                   | Nucleic Acids Res.  |  |          |                        |
| REFERENCE                 | Volume 23 (5), S240 (1998)  |  |          |                        |
| AUTHORS                   | Direct Submission   |  |          |                        |
| TITLE                     | Submitted (11-AUG-1998) Bone & Cartilage Biology, UW 2109,  |  |          |                        |
| JOURNAL                   | SmithKline Beecham, 709 Swedeland Road, King of Prussia, PA 19406, USA  |  |          |                        |
| FEATURES                  | Location/Qualifiers   |  |          |                        |
| source                    | 1..1283   |  |          |                        |
|                           | /organism="Homo sapiens"  |  |          |                        |
|                           | /db_xref="taxon:9606"   |  |          |                        |
|                           | /chromosome="20"  |  |          |                        |
|                           | /map="20q12-q13"  |  |          |                        |
|                           | /cell_type="primary osteoblast"   |  |          |                        |
|                           | 9..761  |  |          |                        |
| CDS                       | /note="CTGF-L; encodes IGF binding (IGFB), Von Willebrand factor type C (WVC) and thrombospondin type I (TSPI) domains: member of the CCN (CTGF/cvrf1/nov) family: lacks        |  |          |                        |

|                           |   |   |              |        |                 |  |  |
|---------------------------|---|---|--------------|--------|-----------------|--|--|
| BASE COUNT                | 216 a   | /db_xref="taxon:9606"   | 418 c        | 390 g  | 242 t           |  |  |
| ORIGIN                    |   |   |              |        |                 |  |  |
| Query Match               | 43.0%   | Score   | 273:         | DB 6:  | Length 1266;    |  |  |
| Best Local Similarity     | 80.9%   | Pred.   | No. 2.5e-55; |        |                 |  |  |
| Matches 318; Conservative | 0;  | Mismatches  | 75;          | Indels | 0; Gaps 0;      |  |  |
| QY                        | 243   | GGTGACATGAGGGCAGCCACCTGATCCATCTTCTTGCGCCACTTCCTTTCCCTCTGCCTTCTC | 302          |        |                 |  |  |
| Db                        | 4   | GGGACATGAGAGGCACACCAGAACCCACCTCTCGGCCCTTCTCCCTCTCTGCCTCCTC      | 63           |        |                 |  |  |
| QY                        | 303   | TCAATGGTGTGCCAGCTGTGCGGACACCCCTGTACCTGTACCTGTGGACACACCCCAG      | 362          |        |                 |  |  |
| Db                        | 64  | TCAAAGTGGGTACCCAGCTGTGCGGACACCATGTACCTGCCCTTGCCACCTGCCCGA       | 123          |        |                 |  |  |
| QY                        | 363   | TGCCCCACAGGGGTACCCCTGGTGTCTGGATGGCTGTGGCTGCTGTAAGTGTGTGCACGG    | 422          |        |                 |  |  |
| Db                        | 124   | TGCCCGCTGGGATGACCCCTGGTGTGGATGGCTGTGGCTGCTGCCGGGTATGTGCACGG     | 181          |        |                 |  |  |
| QY                        | 423   | AGGCTGGGGAGTCTTCGCGACCACTGCATGTCTGCGCACCCACCGCAGGGCCTGGTTTGT    | 482          |        |                 |  |  |
| Db                        | 184   | CGGCTGGGGAGGCCCTGCGACAACCTCCACGTCTGCGACCGCAGGCGCTGTCTGC         | 243          |        |                 |  |  |
| QY                        | 483   | CAGCCTGGGGCAGGCCCTTGGCGCCATGGGGCTGTGTCTTGTGGATGAGGATGACGGT      | 542          |        |                 |  |  |
| Db                        | 244   | CAGCCCGGGCAGGACCCCGTGGCGGGGGCCCTGTGCCTCTTTGGCAGAGGACGACG        | 303          |        |                 |  |  |
| QY                        | 543   | AGCTGTGAGTGAATGGCCGACGCTACCTGGATGGAGAGACCTTTAAACCAATTGCAGG      | 602          |        |                 |  |  |
| Db                        | 304   | AGCTGTGAGTGAACGCGCCCTGTATCGGGAAGGGAGAGACCTTCCAGCCCCACTGCAGC     | 363          |        |                 |  |  |
| QY                        | 603   | GTCCCTGTGCCGTGTGATGACGGTGGCTTCCACC                              | 635          |        |                 |  |  |
| Db                        | 364   | ATCCCTGCCGCTGCGAGGACGGCGCTTCCACC                                | 396          |        |                 |  |  |
| RESULT 8                  |   |   |              |        |                 |  |  |
| AF083500                  |   |   |              |        |                 |  |  |
| LOCUS                     |   |   |              |        |                 |  |  |
| DEFINITION                | Homo sapiens connective tissue growth factor-like protein   | 1283 bp   | mRNA         | linear | PRI 04-NOV-1998 |  |  |
| ACCESSION                 | precursor, mRNA, complete cds.  |   |              |        |                 |  |  |
| VERSION                   | AF083500  |   |              |        |                 |  |  |
| KEYWORDS                  | AF083500.1 GI:3462835   |   |              |        |                 |  |  |
| SOURCE                    | Homo sapiens.   |   |              |        |                 |  |  |
| ORGANISM                  | Homo sapiens  |   |              |        |                 |  |  |
| REFERENCE                 | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.   |   |              |        |                 |  |  |
| AUTHORS                   | Kumar, S., Hand, A.T., Connor, J.C., Dodds, R.A., Ryan, P.J., Trill, J.G., Fisher, S.M., Slemmon, J.R., Lipshutz, D.B., Bartholomew, V., James, I.E., Rieman, D.J., Gowen, M. and Lee, J.C. |   |              |        |                 |  |  |
| TITLE                     | Identification and cloning of CTGF-L from human osteoblasts, a novel cysteine rich protein containing an IGF binding domain   |   |              |        |                 |  |  |
| JOURNAL                   | Nucleic Acids Res.  |   |              |        |                 |  |  |
| REFERENCE                 | Volume 23 (5), S240 (1998)  |   |              |        |                 |  |  |
| AUTHORS                   | Direct Submission   |   |              |        |                 |  |  |
| TITLE                     | Submitted (11-AUG-1998) Bone & Cartilage Biology, UW 2109,  |   |              |        |                 |  |  |
| JOURNAL                   | SmithKline Beecham, 709 Swedeland Road, King of Prussia, PA 19406,  |   |              |        |                 |  |  |
| FEATURES                  | USA   |   |              |        |                 |  |  |
| source                    | Location/Qualifiers   |   |              |        |                 |  |  |
|                           | 1..1283   |   |              |        |                 |  |  |
|                           | /organism="Homo sapiens"  |   |              |        |                 |  |  |
|                           | /db_xref="taxon:9606"   |   |              |        |                 |  |  |
|                           | /chromosome="20"  |   |              |        |                 |  |  |
|                           | /map="20q12-q13"  |   |              |        |                 |  |  |
|                           | /cell_type="primary osteoblast"   |   |              |        |                 |  |  |
|                           | 9..761  |   |              |        |                 |  |  |
| CDS                       | /note="CTGF-L; encodes IGF binding (IGFB), Von Willebrand factor type C (WVC) and thrombospondin type I (TSPI) domains: member of the CCN (CTGF/cvrf1/nov) family: lacks                    |   |              |        |                 |  |  |

the fourth carboxy-terminal domain present in other members of the CCN family"  
/codon\_start=1  
/product="connective tissue growth factor-like protein precursor"  
/protein\_id="AAC70350.1"  
/db\_xref="GI:3462836"  
/translation="MRGTPKTHLLAFSLCLLSKVRTQLCPCTCPWPPRCLGVPLVLDGCGCRVRCARLGPFCDDHVCDSQGLVCPGAGPGGRGALCLLAEEDSSCEVNGRLYREGTEFQPHCSIRCRDGGFTCVPLCSDEVRPLPSWDCPHRVRVEVLGKCCPEWYCGGGGLGTQPLPAQGPQFSLVSLPPGVPCEWSTANGPCSTTCGLGNATRVSNQNRFRLETRQLRLSLRCPSPSRGRSPNSAF"

BASE COUNT 235 a 418 c 389 g 241 t  
ORIGIN

Query Match 43.0%; Score 273; DB 9; Length 1283;

Best Local Similarity 80.9%; Pred. No. 2.5e-55;

Matches 318; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 243 GGTGACATGAGGGGAGCCACCTGATCCATCTTCTGGCCACTTCCTTCTGCTTCCTTC 302  
DB 3 GGGGACATGAGAGGACACCGAAGACCCACCTCTGCGCTTCTGCTTCCTTCCTTCCTTC 62  
QY 303 TCAATGGTGTGCGGAGTGTGCGGACACCTGTACCTGTCTTGGACACACCCAG 362  
DB 63 TCAAAAGGTGCGTACCGAGTGTGCGGACACCATGATCTGCGCCCTGCGCCCTCCCGA 122  
QY 363 TGCCACAGGGGTACCCCTGTGCTGGATGCTGTGGCTGTGTAAGTGTGTCACGG 422  
DB 123 TGCCCGCTGGAGTACCCCTGTGCTGGATGCTGTGGCTGTGCGGATGTGTCACGG 182  
QY 423 AGGCTGGGGAGTCTTGCGACACCTGTGCTGGACCCAGCCAGGCGCTTGTGT 482  
DB 183 CGGCTGGGGAGCCCTTGCGACCACTCCACGCTGCGAGCCAGCGGCTGTGCTGC 242  
QY 483 CAGCTGGGGAGCCCTTGCGGCGCATGGGCTGTGCTCTTGGATGATGACGCT 542  
DB 243 CAGCCCGGGAGGACCCGCTGCGGGGGGCGCTGTGCTCTTGGACAGGACGACG 302  
QY 543 AGCTGTAGGTGAATGGCGGAGTACCTGTGATGAGAGACCTTTAAACCCCAATTCGAGG 602  
DB 303 AGCTGTAGGTGAACGGCGGCTGTATCGGGAAGGGAGACCTTCCAGCCCACTGCAGC 362  
QY 603 GTCTGTGCGCTGTGATGACGGTGTGCTTCAACC 635  
DB 363 ATCCGCTGCGCTGCGAGGACGCGCTTCAACC 395

RESULT 9

AR210322 LOCUS 1293 bp DNA linear PAT 20-JUN-2002  
DEFINITION Sequence 13 from patent US 6387657.  
ACCESSION AR210322  
VERSION AR210322.1 GI:21512523  
KEYWORDS  
SOURCE  
ORGANISM

Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 1293)

AUTHORS Botstein,D.A., Cohen,R.L., Goddard,A.D., Gurney,A.L., Hillan,K.J.,

Lawrence,D.A., Levine,A.J., Pennica,D., Roy,M. Ann. and Wood,W.I.

TITLE WISP polypeptides and nucleic acids encoding same

JOURNAL Patent: US 6387657-A 13 14-MAY-2002;

FEATURES Location/Qualifiers

source 1..1293

/organism="unknown"

BASE COUNT 232 a 425 c 393 g 243 t

ORIGIN

Query Match 43.0%; Score 273; DB 6; Length 1293;

Best Local Similarity 80.9%; Pred. No. 2.5e-55;

Matches 318; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 243 GGTGACATGAGGGGAGCCACCTGATCCATCTTCTGGCCACTTCCTTCTGCTTCCTTC 302  
DB 16 GGGGACATGAGAGGACACCGAAGACCCACCTCTGCGCTTCTGCTTCCTTCCTTCCTTC 75  
QY 303 TCAATGGTGTGCGGAGTGTGCGGACACCTGTACCTGTCTTGGACACACCCAG 362  
DB 76 TCAAAAGGTGCGTACCGAGTGTGCGGACACCATGATCTGCGCCCTGCGCCACTCCCGA 135  
QY 363 TGCCACAGGGGTACCCCTGTGCTGGATGCTGTGGCTGTGTAAGTGTGTCACGG 422  
DB 136 TGCCCGCTGGAGTACCCCTGTGCTGGATGCTGTGGCTGTGCGGCTGTGTCACGG 195  
QY 423 AGGCTGGGGAGTCTTGCGACACCTGTGCTGGACCCAGCGGCTGTGTTGT 482  
DB 196 CGGCTGGGGAGCCCTTGCGACCACTCCACGCTGTGCGAGCCAGCGGCTGTGCTGC 255  
QY 483 CAGCTGGGGAGCCCTTGCGGCGCATGGGCTGTGCTCTTGGATGATGACGCT 542  
DB 256 CAGCCCGGGAGGACCCGCTGCGGCGGCGGCTGTGCTCTTGGACAGGACGACG 315  
QY 543 AGCTGTAGGTGAATGGCGGAGTACCTGTGATGAGAGACCTTTAAACCCCAATTCGAGG 602  
DB 316 AGCTGTAGGTGAACGGCGGCTGTATCGGGAAGGGAGACCTTCCAGCCCACTGCAGC 375  
QY 603 GTCTGTGCGCTGTGATGACGGTGTGCTTCAACC 635  
DB 376 ATCCGCTGCGCTGCGAGGACGCGCTTCAACC 408

RESULT 10

AR210323/c

LOCUS 1293 bp DNA linear PAT 20-JUN-2002

DEFINITION Sequence 14 from patent US 6387657.

ACCESSION AR210323

VERSION AR210323.1 GI:21512524

KEYWORDS

SOURCE

Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 1293)

AUTHORS Botstein,D.A., Cohen,R.L., Goddard,A.D., Gurney,A.L., Hillan,K.J.,

Lawrence,D.A., Levine,A.J., Pennica,D., Roy,M. Ann. and Wood,W.I.

TITLE WISP polypeptides and nucleic acids encoding same

JOURNAL Patent: US 6387657-A 14 14-MAY-2002;

FEATURES Location/Qualifiers

source 1..1298

/organism="unknown"

BASE COUNT 243 a 393 c 425 g 232 t

ORIGIN

Query Match 43.0%; Score 273; DB 6; Length 1293;

Best Local Similarity 80.9%; Pred. No. 2.5e-55;

Matches 318; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 243 GGTGACATGAGGGGAGCCACCTGATCCATCTTCTGGCCACTTCCTTCTGCTTCCTTC 302  
DB 1278 GGGGACATGAGAGGACACCGAAGACCCACCTCTGCGCTTCTGCTTCCTTCCTTCCTTC 1219  
QY 303 TCAATGGTGTGCGGAGTGTGCGGACACCTGTACCTGTCTTGGACACACCCAG 362  
DB 1218 TCAAAAGGTGCGTACCGAGTGTGCGGACACCATGATCTGCGCCCTGCGCCACTCCCGA 1159  
QY 363 TGCCACAGGGGTACCCCTGTGCTGGATGCTGTGGCTGTGTAAGTGTGTCACGG 422  
DB 1158 TGCCCGCTGGAGTACCCCTGTGCTGGATGCTGTGGCTGTGCGGCTGTGTCACGG 1099  
QY 423 AGGCTGGGGAGTCTTGCGACACCTGTGCTGGACCCAGCGGCTGTGTTGT 482  
DB 1098 CGGCTGGGGAGCCCTTGCGACCACTCCACGCTGTGCGAGCCAGCGGCTGTGCTGC 1039  
QY 483 CAGCTGGGGAGCCCTTGCGGCGCATGGGCTGTGCTCTTGGATGATGACGCT 542  
DB 1038 CAGCCCGGGAGGACCCGCTGCGGCGGCGGCTGTGCTCTTGGACAGGACGACG 979



```
Best Local Similarity 80.9%; Pred. No. 2.5e-55;
Matches 318; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 243 GGTGACATGAGGGGAGCCACTGATCCATCTTCTGGCCACTTCTCTCTGCTTCTC 302
Db 150 GGGGACATGAGAGGACACACGAGACCCACTCTCTGGCCCTTCTCTCTGCTTCTC 209
QY 303 TCAATGGTGTGCCCCAGCTGTGCCGACACCTGTACCTGTCTTGGACACCCACG 362
Db 210 TCAAAAGTGCCTACCACTGTGCCGACACCACTGATCTCTGCTTCTCTCTCTC 269
QY 363 TGGCCACAGGGGTACCCCTGTCTGGATGCTGTGGCTGTCTTAAAGTGTGTCAGG 422
Db 270 TCCCGCTGGGAGTACCCCTGTCTGGATGCTGTGGCTGTCTTAAAGTGTGTCAGG 329
QY 423 AGCTGTGGGGAGTCTTGGACACCTGTCTGACCCACCTGTCTGACCCACCTGTTCT 482
Db 330 CGGCTGGGGAGCCCTTGGACCACTTCCACGCTCTGCGAGCCGAGGCTGTGCTGC 389
QY 483 CAGCCTGGGGAGCCCTTGGACCACTTCCACGCTCTGCGAGCCGAGGCTGTGCTGC 542
Db 390 CAGCCGGGGAGCCCTTGGACCACTTCCACGCTCTGCGAGCCGAGGCTGTGCTGC 449
QY 543 AGCTGTGAGTGAATGGCGGAGTACCTGTGATGAGAGACCTTTAAACCAATTCAGG 602
Db 450 AGCTGTGAGTGAATGGCGGAGTACCTGTGATGAGAGACCTTTAAACCAATTCAGG 509
QY 603 GTCCCTGGCGCTGTGATGAGGCTGTTCACC 635
Db 510 ATCCGCTGGCGCTGTGATGAGGCTGTTCACC 542

RESULT 13
BC017782
LOCUS
DEFINITION
Homo sapiens, WNT1 inducible signaling pathway protein 2, clone
MGC:22271 IMAGE:4691574, mRNA, complete cds.
ACCESSION
BC017782
VERSION
BC017782.1 GI:17389482
KEYWORDS
MGC.
SOURCE
Homo sapiens.
ORGANISM
Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1450)
Strausberg, R.
Direct Submission
Submitted (03-DEC-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: CLONTECH
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www.shgc.stanford.edu
Contact: (Dickson, Mark) mcdépxii.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 36 Row: m Column: 3
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 4507922.
Location/Qualifiers
1. .1450
/organism="Homo sapiens"
/db_xref="LocusID:8839"
```

## CDS

```
/db_xref="taxon:9606"
/clone="MGC:22271 IMAGE:4691574"
/tissue_type="Lung"
/clone_lib="NIH_MGC_77"
/lab_host="DH10B"
/note="Vector: pDNR-LIB"
165..917
/codon_start=1
/product="WNT1 inducible signaling pathway protein 2"
/protein_id="AAH17782.1"
/db_xref="GI:17389483"
/translation="MRGTPKTHLLAFSLKLVLTQICPTPCPPPPRCPLGVP
LVLDGGCCRCVRCARLGDHVCDSQGLVCPQPGGRCALCLLAEDDSCEY
NGRLYREGTFTQPHCSIRCEDEGFTVPLCSEDVRLPFWDCPRPRVRLGKCP
WVGOGGGLGTQPLPAOGFQFSGLVSSLPVPCPEWSTAMGPCSTTCLGLMATRVN
QNRFCRLTQRRLCLSRPCPPSRGRSPQNSAF"
BASE COUNT 272 a 457 c 441 g 280 t
ORIGIN
Query Match 43.0%; Score 273; DB 9; Length 1450;
Best Local Similarity 80.9%; Pred. No. 2.5e-55;
Matches 318; Conservative 0; Mismatches 75; Indels 0; Gaps 0;
QY 243 GGTGACATGAGGGGAGCCACTGATCCATCTTCTGGCCACTTCTCTCTGCTTCTC 302
Db 150 GGGGACATGAGAGGACACACGAGACCCACTCTCTGGCCCTTCTCTCTGCTTCTC 218
QY 303 TCAATGGTGTGCCCCAGCTGTGCCGACACCTGTCTGACCCACCTTTCACCCAG 362
Db 210 TCAAAAGTGCCTACCACTGTGCCGACACCACTGATCTCTGCGAGCCGAGGCTGTGCTGC 278
QY 363 TGGCCACAGGGGTACCCCTGTCTGGATGCTGTGGCTGTCTTAAAGTGTGTCAGG 422
Db 270 TCCCGCTGGGAGTACCCCTGTCTGGATGCTGTGGCTGTCTTAAAGTGTGTCAGG 338
QY 423 AGCTGTGGGGAGTCTTGGACACCTGTCTGATGAGAGACCTTTAAACCAATTCAGG 482
Db 330 CGGCTGGGGAGCCCTTGGACCACTTCCACGCTCTGCGAGCCGAGGCTGTGCTGC 398
QY 483 CAGCCTGGGGAGCCCTTGGCGGACCTGTGCTGTCTTAAAGTGTGTCAGG 542
Db 390 CAGCCGGGGAGCCCTTGGCGGACCTGTGCTGTCTTAAAGTGTGTCAGG 458
QY 543 AGCTGTGAGTGAATGGCGGAGTACCTGTGATGAGAGACCTTTAAACCAATTCAGG 602
Db 450 AGCTGTGAGTGAATGGCGGAGTACCTGTGATGAGAGACCTTTAAACCAATTCAGG 518
QY 603 GTCCCTGGCGCTGTGATGAGGCTGTTCACC 635
Db 510 ATCCGCTGGCGCTGTGATGAGGCTGTTCACC 551
```

## RESULT 14

```
AR210338
LOCUS
DEFINITION
Sequence 39 from patent US 6387657.
ACCESSION
AR210338
VERSION
AR210338.1 GI:21512543
KEYWORDS
Unknown.
SOURCE
Unknown.
ORGANISM
Unclassified.
1 (bases 1 to 841)
REFERENCE
1 (bases 1 to 841)
AUTHORS
Botstein, D.A., Cohen, R.L., Goddard, A.D., Gurney, A.L., Hillan, K.J.,
Lawrence, D.A., Levine, A.J., Pennica, D., Roy, M. Ann. and Wood, W.I.
TITLE
WISP polypeptides and nucleic acids encoding same
JOURNAL
Patent: US 6387657-A 39 14-MAY-2002;
FEATURES
Location/Qualifiers
1..841
source
/organism="unknown"
BASE COUNT 124 a 297 c 280 g 140 t
ORIGIN
```







GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 23, 2003, 09:03:08 ; Search time 168.151 Seconds  
(without alignments)  
8504.365 Million cell updates/sec

Title: US-10-010-408-1-copy\_1\_635

Perfect score: 635

Sequence: 1 GAGCGTTCGATCTCCAGAG.....GTGATGACGGTGGCTTCACC 635

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_101002.\*

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Query Match | Score | Length | ID       | Description        |
|------------|-------------|-------|--------|----------|--------------------|
| 1          | 635         | 100.0 | 1708   | AA207516 | Rat HICP polypepti |
| 2          | 521.4       | 82.1  | 1734   | AA207516 | Mouse WISP-2 prote |
| 3          | 387         | 60.9  | 753    | AA207517 | Rat HICP polypepti |
| 4          | 347         | 54.6  | 753    | AA207518 | Mouse WISP-2 prote |
| 5          | 318         | 50.1  | 681    | AA207521 | Rat HICP mature po |
| 6          | 299         | 47.1  | 1522   | AA207521 | Human growth facto |
| 7          | 278.4       | 43.8  | 1337   | AA207521 | Human secreted pro |
| 8          | 278.2       | 43.8  | 1352   | AA207521 | Human secreted pro |
| 9          | 273         | 43.0  | 1266   | AA207521 | Human PRO261 nucle |

|    |       |      |       |    |          |                    |
|----|-------|------|-------|----|----------|--------------------|
| 10 | 273   | 43.0 | 1266  | 22 | AA207516 | Human cDNA sequenc |
| 11 | 273   | 43.0 | 1266  | 22 | AA207516 | PRO261 coding sequ |
| 12 | 273   | 43.0 | 1266  | 22 | AA207516 | Human angiogenesis |
| 13 | 273   | 43.0 | 1267  | 21 | AA207516 | Human PRO261 cDNA  |
| 14 | 273   | 43.0 | 1285  | 19 | AA207516 | Human connective t |
| 15 | 273   | 43.0 | 1293  | 20 | AA207516 | Human WISP-2 prote |
| 16 | 273   | 43.0 | 1309  | 22 | AA207516 | Connective tissue  |
| 17 | 271.4 | 42.7 | 841   | 20 | AA207516 | Human WISP-2 prote |
| 18 | 268.6 | 42.3 | 750   | 20 | AA207516 | Human WISP-2 prote |
| 19 | 268.6 | 42.3 | 1257  | 20 | AA207516 | EGF-like homologue |
| 20 | 261.4 | 41.2 | 738   | 20 | AA207516 | Human WISP-2 prote |
| 21 | 210   | 33.1 | 210   | 20 | AA207516 | Rat HICP IGFBP dom |
| 22 | 203.4 | 32.0 | 2136  | 22 | AA207516 | Human full-length  |
| 23 | 164   | 25.8 | 13255 | 22 | AA207516 | Human immune/haema |
| 24 | 122   | 19.2 | 634   | 22 | AA207516 | Human cDNA 5'-end  |
| 25 | 122   | 19.2 | 634   | 22 | AA207516 | Human cDNA clone r |
| 26 | 90    | 14.2 | 177   | 20 | AA207516 | Rat HICP VNC domai |
| 27 | 87.8  | 13.8 | 1062  | 20 | AA207516 | Human connective t |
| 28 | 87.8  | 13.8 | 1062  | 24 | ABK88299 | Human connective t |
| 29 | 85.6  | 13.5 | 1146  | 18 | AA207516 | Human monocyte mat |
| 30 | 85.6  | 13.5 | 1146  | 24 | ABK88299 | Human connective t |
| 31 | 85.6  | 13.5 | 1418  | 24 | ABK88299 | DNA encoding human |
| 32 | 85.6  | 13.5 | 1419  | 24 | ABK88299 | Human Cyr61 protei |
| 33 | 85.6  | 13.5 | 1887  | 24 | ABK88299 | Human osteoblast d |
| 34 | 85.6  | 13.5 | 2016  | 22 | AA207516 | Human shear stress |
| 35 | 85.6  | 13.5 | 2016  | 24 | ABK88299 | Human osteoblast d |
| 36 | 85.6  | 13.5 | 2016  | 24 | ABK88299 | Human benign prost |
| 37 | 85.6  | 13.5 | 2016  | 24 | ABK88299 | Kidney cancer rela |
| 38 | 85.6  | 13.5 | 2016  | 24 | ABK88299 | Human Cyr61 protei |
| 39 | 85.6  | 13.5 | 2021  | 24 | ABK88299 | Human osteoblast d |
| 40 | 85.6  | 13.5 | 2025  | 24 | ABK88299 | Human osteoblast d |
| 41 | 85.6  | 13.5 | 2052  | 24 | ABK88299 | Human osteoblast d |
| 42 | 85.6  | 13.5 | 2307  | 21 | AA207516 | Human cancer assoc |
| 43 | 84    | 13.2 | 1418  | 18 | AA207516 | Human cysteine ric |
| 44 | 84    | 13.2 | 1418  | 22 | AA207516 | Human cysteine-ric |
| 45 | 83.8  | 13.2 | 1480  | 22 | AA207516 | Mouse cysteine-ric |

#### ALIGNMENTS

RESULT 1  
AA207516  
ID AA207516 standard; cDNA; 1708 BP.  
XX AC AA207516;  
XX DT 26-NOV-1999 (first entry)  
XX DE Rat HICP polypeptide encoding cDNA.  
XX KW Heparin-induced CCN-like protein; HICP; cell-associated activity; ss;  
XX KW cardiovascular disorder; aberrant cell proliferation; fibrotic disorder.  
XX OS Rattus sp.  
XX PN WO9947556-A2.  
XX PD 23-SEP-1999.  
XX PF 18-MAR-1999; 99WO-US059999.  
XX PR 19-MAR-1998; 98US-0044273.  
XX PA (TUFT ) TUFTS COLLEGE.  
XX PI Castellot JJ;  
XX WPI: 1999-562060/47.  
XX P-PSDB; AA27434.  
XX PT Nucleic acid sequences encoding rat heparin-induced CCN-like protein, used in methods to identify modulators or in diagnostic applications

XX PS Claim 2; Fig 1; 108pp; English.

XX This cDNA encodes a rat heparin-induced CCN-like protein (HICP) protein.

CC Agents that stimulate or inhibit HICP protein activity or expression,

CC antisense HICP nucleic acid molecules and HICP antibodies, can be used to

CC modulate cell-associated activity. HICP modulators can be used to treat

CC disorders characterized by aberrant HICP protein activity or expression.

CC Probes capable of hybridizing to HICP mRNA or antibodies specific for

CC HICP can be used to detect HICP activity in a biological sample. HICP

CC can be used to treat disorders, such as a cardiovascular or fibrotic

XX disorder, characterized by aberrant cell proliferation.

XX Sequence 1708 BP; 362 A; 486 C; 478 G; 382 T; 0 other;

Query Match 100.0%; Score 635; DB 20; Length 1708;

Best Local Similarity 100.0%; Pred. No. 1.9e-170;

Matches 635; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGCTTCTGATCTCCAGAGACCCCTGGGGTGGGACAGGGGCTTGGCAAGGCTCGAGCC 60

DB 1 GACGCTTCTGATCTCCAGAGACCCCTGGGGTGGGACAGGGGCTTGGCAAGGCTCGAGCC 60

QY 61 GCTGGCAGTGGCTTGGATGAGGCTCTTATTACTGGAACTGAGGAGCTTAAGAGGCTC 120

DB 61 GCTGGCAGTGGCTTGGATGAGGCTCTTATTACTGGAACTGAGGAGCTTAAGAGGCTC 120

QY 121 CTGTGAGCTTGTCTTAAAGTCTTAGCACTTGTGTGGCTTGGGCTTCACACACTCTCAGA 180

DB 121 CTGTGAGCTTGTCTTAAAGTCTTAGCACTTGTGTGGCTTGGGCTTCACACACTCTCAGA 180

QY 181 CACCTTGTGTGGCTTCCAGGCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCA 240

DB 181 CACCTTGTGTGGCTTCCAGGCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCA 240

QY 241 ACGGTGACATGAGGGGACCCACATGATCCATCTTCTGGCCACTTCTTCTCTCTCTCTCT 300

DB 241 ACGGTGACATGAGGGGACCCACATGATCCATCTTCTGGCCACTTCTTCTCTCTCTCTCTCT 300

QY 301 TCTCAATGTGTGTGCTTCCAGGCTTGTGCGGACACCCCTGTACCTGTCTTGTGACACACCC 360

DB 301 TCTCAATGTGTGTGCTTCCAGGCTTGTGCGGACACCCCTGTACCTGTCTTGTGACACACCC 360

QY 361 AGTGGCCACAGGGGTACCCCTGTGCTGGATGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 420

DB 361 AGTGGCCACAGGGGTACCCCTGTGCTGGATGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 420

QY 421 GGAGCTGGGGGAGTCTTGGACACCTGATGCTGTGCGACCCAGCCAGGCGCTTGGTT 480

DB 421 GGAGCTGGGGGAGTCTTGGACACCTGATGCTGTGCGACCCAGCCAGGCGCTTGGTT 480

QY 481 GTACGCTGGGGGAGGCTTGGGCGCCATGGGGCTGTGTGTCTTGTGATGAGGATGAGC 540

DB 481 GTACGCTGGGGGAGGCTTGGGCGCCATGGGGCTGTGTGTCTTGTGATGAGGATGAGC 540

QY 541 GTAGCTGTGAGTGAATGGCCGACGCTACCTGGATGGAGACCTTAAACCAATTGCA 600

DB 541 GTAGCTGTGAGTGAATGGCCGACGCTACCTGGATGGAGACCTTAAACCAATTGCA 600

QY 601 GGGTCTGTGGCGCTGTGATGACGCTGCTTCAAC 635

DB 601 GGGTCTGTGGCGCTGTGATGACGCTGCTTCAAC 635

RESULT 2

AA76488

ID AA76488 standard; DNA; 1734 BP.

XX

AC AA76488;

XX

XX 06-AUG-1999 (first entry)

XX

DE Mouse WISP-2 protein nucleotide sequence SEQ ID NO:17.

XX WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;

KW connective tissue growth factor; cancer; melanoma; arteriosclerosis;

KW leukaemia; lymphoid malignancy; haematopoiesis-related disorder;

KW tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;

KW kidney disorder; bone-related disorder; osteoporosis; trauma; burn;

KW connective tissue disorder; catabolic state; inflammation;

KW testicular-related disorder; angiogenesis; immunological disorder; ss.

XX Mus sp.

XX WO9921998-A1.

PN 06-MAY-1999.

XX 29-OCT-1998; 98WO-US22991.

PF 14-APR-1998; 98US-0081695.

PR 29-OCT-1997; 97US-0063704.

PR 03-FEB-1998; 98US-0093612.

XX (GETH ) GENENTECH INC.

XX Botstein DA, Cohen RL, Goddard A, Gurney AL, Hillan K;

PI Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WI;

DR WPI; 1999-337420/28.

DR P-PSDB; AAY17651.

PT New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3

XX Example 2; Page 178-179; 28pp; English.

XX The present invention describes Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2 and WISP-3 have homology to connective tissue growth factor (CTGF). Products from the present invention can be used to treat WISP-related disorders such as breast, ovarian, and colon cancer or melanoma. The products can be used to treat arteriosclerosis. The products can also be used to treat other diseases e.g. benign and malignant tumours, leukaemia and lymphoid malignancies, neuronal, glial, astrocytal, hypothalamic and other glandular, macrophage, epithelial, stromal, and blastocoele disorders, haematopoiesis-related disorders, tissue-growth disorders, skin disorders, desmoplasia, fibrotic lesions, kidney disorders, bone-related disorders such as osteoporosis, trauma such as burns, incisions, and other wounds, connective tissue disorders, catabolic states, testicular-related disorders, and inflammatory, angiogenic and immunologic disorders including arteriosclerosis. The products can also be used for detection and diagnosis especially of individuals with neoplastic cell growth or proliferation. The products can be used in the production of transgenic or knock-out animals. Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing cells.

XX Sequence 1734 BP; 355 A; 491 C; 495 G; 393 T; 0 other;

QY 3 CGCTTCGATCTCCAGAGGACCCCTGGGTGGGACAGGGGCTTGGCAAGGCTGACGCGC 62

DB 13 CGCTTCGATCTCCAGAGGACCCGCGGTGGGACAGGGGCTTGGCGAGGCTGACGCGC 72

QY 63 TG-GGCAGTGGCTTGGAAATGGAGGTCTTTATTACTGGGAACCTAGAGAGCTAAGAGCTCC 121

DB 73 TGTGGCAGTACCTTGGGATGGAGGTCTTTCTTGTGGGAACCTAGAGAGCTGACGCTCC 132

QY 122 TGTGAG---CTGTGCTTAAGTCTTAGCACTTGTGTGGCTTGGGCTTTCACACACTGTCA 178

DB 133 TGTGAGGCTCTGTCTTAAGTCTTAGCACTTGTGGCCTTGGGCTTTCACACACTGTCA 192

QY 179 GACACCTTGTGGTGGGCTTCCAGGCGCTCACCTTGTGAGCTGGCTCCACAGGG 238

Query Match 82.1%; Score 521.4; DB 20; Length 1734;

Best Local Similarity 92.0%; Pred. No. 3.8e-138;

Matches 586; Conservative 0; Mismatches 41; Indels 10; Gaps 3;











OY 611 CCGCTGTGATGACGGTGGCTTCACC 635  
||||| || ||||| ||||| |||||  
Db 415 CCGCTGGGAGGACGGCGCTTCACC 439

## RESULT 9

AAA30048  
ID AAA30048 standard; cDNA; 1266 BP.

XX  
AC AAA30048;

XX  
DT 09-AUG-2000 (first entry)

XX  
DE Human PRO261 nucleotide sequence.

XX  
KW Antisense; PRO187; PRO533; PRO214; PRO240; PRO211; PRO230; PRO261; PRO246;  
PRO317; tumour growth inhibitor; cancer; diagnosis; treatment; human;  
cell growth; proliferation; growth factor; ADEPT;  
KW antibody dependent enzyme mediated prodrug therapy; ss.

XX  
OS Homo sapiens.

XX  
PN WO200015666-A2.

XX  
PD 23-MAR-2000.

XX  
PF 08-SEP-1999; 99WO-US20594.

XX  
PR 10-SEP-1998; 98US-009803.

XX  
PR 10-SEP-1998; 98WO-US18824.

XX  
PA (GETH ) GENENTECH INC.

XX  
PI Goddard A, Gurney AL, Hillan KJ, Roy MA, Wood WI, Botstein D;

XX  
XX WPI; 2000-271386/23.

DR  
P-PSDB; AAY88573.

XX  
PT New isolated antibodies which bind to specific polypeptides used for  
diagnosis and treatment of neoplastic cell growth and proliferation -

XX  
PS Example 7; Fig 13; 200pp; English.

XX  
CC This sequence represents a human PRO261 nucleotide sequence. PRO261 is a  
growth factor. The invention relates to isolated antibodies which bind to  
a polypeptide. The "PRO" polypeptides are encoded by genes which are over  
expressed in the genome of tumour cells. Vectors and host cells  
comprising the nucleic acid encoding the antibodies are used in the  
production of the antibodies. The antibodies and nucleic acids encoding  
them are used for diagnosing a tumour in a mammal. The antibodies are  
used for inhibiting the growth of tumour cells and identifying compounds  
that inhibit a biological or immunological activity of and/or expression  
of a PRO187, PRO533, PRO214, PRO240, PRO211, PRO230, PRO261, PRO246 or  
PRO317 polypeptide. The antibody can be used in antibody dependent enzyme  
mediated prodrug therapy (ADEPT) by conjugating the antibody to a  
prodrug-activating enzyme which converts a prodrug to an anti-cancer  
drug. The antibodies can be fluorescently labelled and monitored by light  
microscopy, flow cytometry or fluorimetry for diagnosis and prognosis of  
tumours.

XX  
SQ Sequence 1266 BP; 216 A; 418 C; 390 G; 242 T; 0 other;

Query Match 43.0%; Score 273; DB 21; Length 1266;

Best Local Similarity 80.9%; Pred. No. 1.4e-67;

Matches 318; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

OY 243 GGTGACATGAGGCGACCCACTGATCATCTTCTGGCCACTTCTTCTCTGCTCTCTC 302  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 4 GGGGACATGAGGCGACACCCGACCCACTTCTGGCTTCTCTCTCTCTCTCTCTC 63

OY 303 TCAATGGTGTGCGCCAGCTGTGGGACACCCCTGTACCTGTCTTGGACACACCCGAG 362  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 64 TCAAAGGTGCGTACCCAGCTGTGGCCGACACCATGTACCTGCCCTGCCACCTCCCGA 123

OY 363 TGCCACAGGGGTACCCCTGGTGTGGATGGCTGTGGCTGTCTGTAAAGTGTGTGCACGG 422  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 124 TGCCCGCTGGGAGTACCCCTGGTGTGGATGGCTGTGGCTGTCTGTCCGGGTATGTGCACGG 183  
OY 423 AGGCTGGGGAGTCTTCGACCCACTTGCATGTGTGCGACCCGACCCAGGGGCTGTGTGTGT 482  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 184 CGGCTGGGGAGGCTTCGACCCACTTGCATGTGTGCGACCCGACCCAGGGGCTGTGTGTGT 243  
OY 483 CAGCCTGGGCGAGCCCTGGGCGCATGGGCGCTGTGTCTCTTTGGATGAGGATGACGGT 542  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 244 CAGCCCGGGGAGGACCCCGTGGCGGGGGGCCCTGTGCTCTTTGGCAGAGGACGACGC 303  
OY 543 AGCTGTGAGGTGAATGCGCGAGGTACTCTGATGGAGAGACCTTTAAACCCCAATTGCAGG 602  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 304 AGCTGTGAGGTGAACGCGCGCTGTATCGGGAAGGGGAGACCTTTCCAGCCCACTGCAGC 363  
OY 603 GTCTGTGCGCGCTGTGATGACGGTGGCTTCACC 635  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 364 ATCCGCTGCGGCTGCGAGGACGGCGCTTCACC 396

## RESULT 10

AAS21403

ID AAS21403 standard; cDNA; 1266 BP.

XX  
AC AAS21403;

XX  
DT 24-OCT-2001 (first entry)

XX  
DE Human cDNA sequence encoding for PRO261 polypeptide.

XX  
KW Human secretory and transmembrane; PRO; mammalian; cancer; lung;  
breast; prostate; cervical; tumour necrosis factor-alpha; TNF-alpha;  
cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;  
adipocyte; A-peptide; factor VIIA; gene therapy; ss.

XX  
OS Homo sapiens.

XX  
PN WO200140466-A2.

XX  
PD 07-JUN-2001.

XX  
PF 01-DEC-2000; 2000WO-US32678.

XX  
PR 01-DEC-1999; 99WO-US28301.

XX  
PR 01-DEC-1999; 99WO-US28634.

XX  
PR 02-DEC-1999; 99WO-US28551.

XX  
PR 02-DEC-1999; 99WO-US28564.

XX  
PR 02-DEC-1999; 99WO-US28565.

XX  
PR 09-DEC-1999; 99US-0170262.

XX  
PR 16-DEC-1999; 99WO-US30095.

XX  
PR 20-DEC-1999; 99WO-US30911.

XX  
PR 20-DEC-1999; 99WO-US30999.

XX  
PR 30-DEC-1999; 99WO-US31243.

XX  
PR 06-JAN-2000; 2000WO-US00277.

XX  
PR 11-FEB-2000; 2000WO-US0376.

XX  
PR 18-FEB-2000; 2000WO-US03565.

XX  
PR 18-FEB-2000; 2000WO-US04341.

XX  
PR 22-FEB-2000; 2000WO-US04342.

XX  
PR 24-FEB-2000; 2000WO-US04914.

XX  
PR 24-FEB-2000; 2000WO-US05004.

XX  
PR 01-MAR-2000; 2000WO-US05601.

XX  
PR 20-MAR-2000; 2000WO-US07377.

XX  
PR 21-MAR-2000; 2000WO-US07532.

XX  
PR 30-MAR-2000; 2000WO-US08439.

XX  
PR 17-MAY-2000; 2000WO-US13705.

XX  
PR 22-MAY-2000; 2000WO-US14042.

XX  
PR 30-MAY-2000; 2000WO-US14941.

XX  
PR 02-JUN-2000; 2000WO-US15264.

XX  
PR 10-NOV-2000; 2000WO-US30873.

```
PA (GETH ) GENENTECH INC.
XX Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;
PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
XX WPI; 2001-408281/43.
DR P-PSDB; AAU12331.
XX
XX Isolated, secretory and transmembrane PRO polypeptide used to detect
PT other PRO polypeptides, link bioactive molecules to cells expressing
PT PRO polypeptides, and detect the presence of mammalian tumours e.g.
PT lung, breast, prostate, cervical
XX
XX Claim 3; Fig 319; 813pp; English.
XX
XX AAS21244-AAS21518 encode for novel human secretory and transmembrane
CC PRO polypeptides. The PRO polypeptides are useful to detect other
CC PRO polypeptides, to link bioactive molecules to cells expressing
CC PRO polypeptides, to modulate biological activities of cells expressing
CC PRO polypeptides, and to detect the presence of mammalian lung, colon,
CC breast, prostate, rectal, cervical or liver tumours by comparing PRO
CC polypeptide expression in a cell sample to that in a control sample.
CC Some of the 275 sequences are also useful to stimulate the release of
CC tumour necrosis factor-alpha (TNF-alpha) from human blood, the
CC proliferation or differentiation of chondrocytes, the proliferation or
CC gene expression in pericyte cells, the release of proteoglycans from
CC cartilage, the proliferation of inner ear utricular supporting cells or
CC of T-lymphocytes, the release of a cytokine from peripheral blood
CC monocytes (PBMCs), or the proliferation of endothelial cells. Some of
CC the PRO polypeptides may modulate glucose or free fatty acid uptake by
CC skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide
CC to factor VIIA. The PRO polypeptides can be used in assays to identify
CC molecules involved in binding interactions. The polynucleotides encoding
CC PRO polypeptides can be used to generate probes, antisense RNA/DNA,
CC transgenic or knock out animals and can be used in gene therapy.
XX
XX Sequence 1266 BP; 216 A; 418 C; 390 G; 242 T; 0 other;

Query Match 43.0%; Score 273; DB 22; Length 1266;
Best Local Similarity 80.9%; Pred. No. 1.4e-67;
Matches 318; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 243 GGTGACATGAGGGGAGCCGACCTGATCCATCTTCCTGGCCACTTCCTCTGCTGCTTC 302
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 303 TCAATGGTGTGCGGACCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 362
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 363 TGCCACACAGGGGTTACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 422
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 423 AGGCTGGGGAGTCTGCGACCACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 482
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 483 CAGCTGGGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 542
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 543 AGCTGTAGGTGAATGGCGGAGGTACCTGGATGGAGAGACCTTTTAAACCCCAATTTGACGG 602
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 603 GTCTGTGCGCGTGTGATGACGGTGGCTTCAACC 635
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 636 ATCCGCTGCGCGTGGCGAGGACGGCGGCTTCAACC 396
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 11
```

AAF60368

ID AAF60368 standard; cDNA; 1266 BP.

XX

AC AAF60368;

XX

DT 27-APR-2001 (first entry)

XX

DE PRO261 coding sequence.

XX

KW Cytostatic; PRO protein; tumour; cancer; ss.

XX

OS Homo sapiens.

XX

PN WO200105836-A1.

XX

PD 25-JAN-2001.

XX

PF 20-DEC-1999; 99WO-US80999.

XX

PR 20-JUL-1999; 99US-0144758.

XX

PR 26-JUL-1999; 99US-0145698.

XX

PR 08-SEP-1999; 99WO-US20594.

XX

PR 13-SEP-1999; 99WO-US20944.

XX

PR 15-SEP-1999; 99WO-US21090.

XX

PR 05-OCT-1999; 99WO-US23089.

XX

PR 29-NOV-1999; 99WO-US28214.

XX

PR 30-NOV-1999; 99WO-US28313.

XX

PR 02-DEC-1999; 99WO-US28564.

XX

PA (GETH ) GENENTECH INC.

XX

PI Botstein D, Goddard A, Gurney AL, Hillan KJ, Roy MA, Wood WI;

XX

WPI; 2001-091968/10.

XX

DR P-PSDB; AAB68598.

XX

PT New antibody that binds to a PRO polypeptide, e.g. PRO187 and PRO533,

XX

PT useful for diagnosing and treating cancers -

XX

Claim 50; Fig 13; 196pp; English.

XX

PS The present invention relates to PRO proteins and coding sequences. The

XX

CC present sequence is the coding sequence for one such PRO protein.

XX

CC It was found that the PRO genes are amplified in the genome of tumour

XX

CC cells. The gene amplification is expected to be associated with the

XX

CC overexpression of the gene product and contributes to tumorigenesis.

XX

CC Therefore, antagonists of PRO proteins are useful for the treatment of

XX

CC benign or malignant tumours, leukaemias, lymphoid malignancies and other

XX

CC disorders such as neuronal, glial, astrocytal, hypothalamic, glandular,

XX

CC epithelial, inflammatory and immunologic disorders.

XX

SQ Sequence 1266 BP; 216 A; 418 C; 390 G; 242 T; 0 other;

XX

Query Match 43.0%; Score 273; DB 22; Length 1266;

Best Local Similarity 80.9%; Pred. No. 1.4e-67;

Matches 318; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 243 GGTGACATGAGGGGAGCCGACCTGATCCATCTTCCTGGCCACTTCCTCTGCTGCTTC 302

DB 4 GGGGACATGAGAGGACACACCCGAGACCCACCTCTCTGCTGCTTCCTCTGCTGCTTC 63

QY 303 TCAATGGTGTGCGGACCTGTGCGGACACCCCTGTACCTGTCTTGGACACACCCACG 362

DB 64 TCAAAAGGTGCGTAGCAGCTGTGCGGACACCATGTACCTGCGCCCTGGCCACCTCCCGA 123

QY 363 TGCCACACAGGGGTTACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 422

DB 124 TGCCCGCTGGGAGTACCCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 482

QY 423 AGGCTGGGGAGTCTGCGACCACTGATGCTGCGACCCCGACCCAGCGGCTGGTTGT 482

DB 184 CGGCTGGGGAGGCTTGGGACCACTCCACGCTCTCGGAGCCGACGCGGCTGGTCTGC 243

QY 483 CAGCCTGGGAGAGCCCTGGCGCCATGAGGGCTGTGTCTTGTGGATGAGGATGACGGT 542  
 DB 244 CAGCCCGGGGAGACACCGGTGGCGGGGGCCCTGTGCTCTTGGCAGAGACGACAGC 303  
 QY 543 AGCTGTGAGTGAATGCGCCAGCTACCTGATGGAGAGACCTTAAACCCCAATTGCAGG 602  
 DB 304 AGCTGTGAGTGAACGGCCGCTGTATCGGAAGGGGAGACCTTCCAGCCCCACTGCAGC 363  
 QY 603 GTCTGTGCGCTGTGATGACGGTGGCTTACC 635  
 DB 364 ATCCGCTGCCCTGCGAGGACGGCGGCTTACC 396

RESULT 12  
 AAC97451  
 ID AAC97451 standard; cDNA; 1266 BP.  
 XX AC AAC97451;  
 XX DT 28-FEB-2001 (first entry)  
 XX DE Human angiogenesis-associated protein PRO261 cDNA, SEQ ID NO:107.  
 XX KW Human; angiogenesis-associated protein; PRO; endothelial cell growth;  
 KW cardiac hypertrophy; cardiovascular disorder; endothelial disorder;  
 KW angiogenic disorder; atherosclerosis; osteoporosis; hypertension;  
 KW myocardial infarction; diabetic retinopathy; rheumatoid arthritis;  
 KW Crohn's disease; psoriasis; endometriosis; ulcer; wound healing; cancer;  
 KW Alzheimer's disease; Huntington's disease; stroke; drug screening;  
 KW gene therapy; transgenic animal; ss.  
 XX OS Homo sapiens.  
 XX PN W0200053753-A2.  
 XX PD 14-SEP-2000.  
 XX PF 05-JAN-2000; 2000WO-US00219.  
 XX PR 08-MAR-1999; 99WO-US05028.  
 PR 12-MAR-1999; 99US-0123957.  
 PR 14-MAY-1999; 99US-0134287.  
 PR 02-JUN-1999; 99WO-US12252.  
 PR 23-JUN-1999; 99US-0141037.  
 PR 20-JUL-1999; 99US-0144758.  
 PR 26-JUL-1999; 99US-0145698.  
 PR 01-SEP-1999; 99WO-US20111.  
 PR 08-SEP-1999; 99WO-US20594.  
 PR 15-SEP-1999; 99WO-US21090.  
 PR 15-SEP-1999; 99WO-US21547.  
 PR 05-OCT-1999; 99WO-US23089.  
 PR 30-NOV-1999; 99WO-US28313.  
 PR 30-NOV-1999; 99WO-US28409.  
 PR 02-DEC-1999; 99WO-US28564.  
 PR 02-DEC-1999; 99WO-US28565.  
 XX (GETH ) GENENTECH INC.  
 XX PA Ashkenazi AJ, Baker KP, Ferrara N, Gerber H, Goddard A;  
 PI Godowski PJ, Gurney AL, Hillan KJ, Kuo SS, Mark MR, Marsters SA;  
 PI Piloni NF, Pitti RM, Watanabe CK, Williams PM, Wood WI;  
 XX WPI; 2001-090793/10.  
 DR P-PSDB; AAB53084.  
 XX  
 XX New isolated nucleic acid for producing a PRO polypeptide, analyzing  
 PT genetic disorders and treating cardiovascular, endothelial or  
 PT angiogenic disorders, such as atherosclerosis, wounds or cancer -  
 XX  
 XX Claim 58; Fig 41; 293pp; English.  
 XX  
 XX The invention relates to novel human angiogenesis-associated proteins  
 CC designated PRO proteins (AAB53064-B53097), and to nucleic acids encoding

CC PRO proteins. The invention also relates to vectors and host cells  
 CC comprising a PRO nucleic acid, the recombinant production of a PRO  
 CC protein, PRO antibodies specific for a PRO protein, fusion proteins  
 CC comprising a PRO protein, agonists or antagonists of a PRO protein, and  
 CC compounds which inhibit the expression of a PRO gene. The invention  
 CC additionally encompasses methods of identifying modulators of PRO  
 CC expression or activity; diagnosing a cardiovascular, endothelial or  
 CC angiogenic disorder, or a susceptibility to such a disorder by detecting  
 CC mutations in a PRO gene, or the expression level of a PRO gene within a  
 CC particular tissue; treating a cardiovascular, endothelial or angiogenic  
 CC disorder via the administration of a PRO protein, PRO nucleic acid, or  
 CC PRO agonist or antagonist; a retroviral gene therapy vector comprising a  
 CC PRO nucleic acid; and methods of inhibiting or stimulating endothelial  
 CC cell growth, cardiac hypertrophy or PRO-induced angiogenesis via the  
 CC administration of a PRO protein, or an agonist or antagonist thereof.  
 CC PRO nucleic acids, PRO proteins, antibodies against PRO proteins, PRO  
 CC agonists and PRO antagonists may be used as therapeutic agents to treat  
 CC cardiovascular, endothelial or angiogenic disorders, such as  
 CC atherosclerosis, osteoporosis, myocardial infarction, hypertension,  
 CC diabetic retinopathy, rheumatoid arthritis, Crohn's disease, psoriasis,  
 CC endometriosis, ulcers, wounds, cancer, Alzheimer's disease, Huntington's  
 CC disease, or stroke. PRO nucleic acids are additionally useful in the  
 CC recombinant production of PRO proteins, as hybridisation probes to  
 CC screen libraries to isolate cDNAs with sequence identity to PRO proteins,  
 CC to map genes encoding PRO proteins, to analyse genetic disorders, and in  
 CC gene therapy. PRO nucleic acids can also be used to produce transgenic  
 CC animals useful for the development and screening of potential  
 CC therapeutic agents. The present sequence represents a cDNA encoding a PRO  
 CC protein of the invention.

XX SQ Sequence 1266 BP; 216 A; 418 C; 390 G; 242 T; 0 other;

Query Match 43.0%; Score 273; DB 22; Length 1266;  
 Best Local Similarity 80.9%; Pred. No. 1.4e-67;  
 Matches 318; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 243 GGTGACATGAGGGGACGCCACCTGATCCATCTTCTGGCCACTTCTCTTCTGCTGCTC 302  
 DB 4 GGGGACATGAGAGGACACCGAAGACCCACTCTCTGGCCCTTCTCCCTCTCTGCTCCTC 63  
 QY 303 TCAATGCTGTGCCAGCTGTGCCGGACACCCCTGTACCTGTCTTGGACACACCCAG 362  
 DB 64 TCAAAAGTGGCTACCCAGCTGTGCCCGACACCATGTACCTGCCCTGCCACCTCCCGGA 123  
 QY 363 TGCCACACAGGGGTACCCCTGGTGGCTGGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 422  
 DB 124 TGCCCGCTGGAGGTACCCCTGGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTG 183  
 QY 423 AGGCTGGGGAGTCTTCGACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 482  
 DB 184 CGGCTGGGGAGGCTCGGACCACTCCACGTCTGCGACGCCAGCCAGGGGCTGCTGCTGCTG 243  
 QY 483 CAGCCTGGGGAGGCTCGGCGCCATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 542  
 DB 244 CAGCCCGGGAGGACCCCGTGGCGGGGGCCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 303  
 QY 543 AGCTGTGAGTGAATGCGCCAGGTACCTGGATGGAGAGACCTTTAAACCAATTCGAGG 602  
 DB 304 AGCTGTGAGTGAACGGCCGCCCTGTATCGGAAGGGGAGAGACCTTCCAGCCCCACTGCAGC 363  
 QY 603 GTCTGTGCGCTGTGATGACGGTGGCTTACC 635  
 DB 364 ATCCGCTGCCCTGCGAGGACGGCGGCTTACC 396

RESULT 13  
 AAA77566  
 ID AAA77566 standard; cDNA; 1267 BP.  
 XX AC AAA77566;  
 XX DT 07-NOV-2000 (first entry)  
 XX

Human PRO261 cDNA sequence SEQ ID NO:71.

DE Human; PRO; promotion; inhibition; angiogenesis; cardiovascularisation;  
KW diagnosis; trauma; wound; cancer; atherosclerosis; cardiac hypertrophy;  
KW angiogenic; proliferative; cardiac; cardiovascular; antiatherosclerotic;  
KW cytostatic; gene therapy; vaccine; ss.  
XX Homo sapiens.  
OS  
XX  
PN WO200032221-A2.  
XX  
XX 08-JUN-2000.  
XX  
PF 30-NOV-1999; 99WO-US28313.  
XX  
PR 01-DEC-1998; 98WO-US25108.  
PR 16-DEC-1998; 98US-0112850.  
PR 12-JAN-1999; 99US-0115554.  
PR 08-MAR-1999; 99WO-US05028.  
PR 12-MAR-1999; 99US-0123957.  
PR 28-APR-1999; 99US-0131445.  
PR 14-MAY-1999; 99US-0134287.  
PR 02-JUN-1999; 99WO-US12252.  
PR 23-JUN-1999; 99US-0141037.  
PR 20-JUL-1999; 99US-0144758.  
PR 26-JUL-1999; 99US-0145698.  
PR 01-SEP-1999; 99WO-US20111.  
PR 08-SEP-1999; 99WO-US20594.  
PR 13-SEP-1999; 99WO-US20944.  
PR 15-SEP-1999; 99WO-US21090.  
PR 15-SEP-1999; 99WO-US21547.  
PR 05-OCT-1999; 99WO-US23089.  
PR 29-OCT-1999; 99US-0162506.  
XX  
XX (GETH ) GENENTECH INC.  
XX  
PI Ashkenazi AJ, Baker KP, Ferrara N, Gerber H, Hillan KJ, Goddard A;  
PI Godowski PU, Gurney AL, Klein RD, Kuo SS, Paoni NF, Smith V;  
PI Watanabe CK, Williams PM, Wood WI;  
XX  
DR WPI; 2000-412154/35.  
DR P-PSDB; AAB24402.  
XX  
XX Nucleic acids encoding PRO polypeptides useful for preventing,  
PT diagnosing and treating a cardiovascular, endothelial or  
PT angiogenic disorders in mammals -  
XX  
PS Claim 61; Fig 29; 315pp; English.  
XX  
XX The present invention describes nucleic acids encoding PRO polypeptides  
CC useful for preventing, diagnosing and treating a  
CC cardiovascular, endothelial or angiogenic disorder in mammals by  
CC modulating cell proliferation, angiogenesis and cardiovascularisation,  
CC and for identifying agonists and antagonists of these processes. The  
CC nucleic acids and the proteins they encode may be used in the  
CC prevention, treatment and diagnosis of diseases associated with  
CC inappropriate PRO expression such as cardiovascular, endothelial or  
CC angiogenic disorders in mammals (e.g. atherosclerosis, cancers and  
CC cardiac hypertrophy). For example, the nucleic acids (NCs) and vectors  
CC containing them and the PRO polypeptide may be used to treat disorders  
CC associated with decreased PRO expression. AAA77510 to AAA77721 and  
CC AAB24388 to AAB24435 represent nucleotide and protein sequences used in  
CC the exemplification of the present invention.  
XX  
SQ Sequence 1267 BP; 216 A; 419 C; 390 G; 242 T; 0 other;  
XX  
Query Match 43.0%; Score 273; DB 21; Length 1267;  
Best Local Similarity 80.9%; Pred. No. 1.4e-67;  
Matches 318; Conservative 0; Mismatches 75; Indels 0; Gaps 0;  
QY 243 GGTGACATGAGGGGAGCCCACTGATCCATCTCTGCGCACTCTCTCTGCTTCCTC 302  
DB ||||||||| ||| || ||||| ||||||| ||||||||| |||  
4 GGGGACATGAGGGGAGCCCACTGATCCATCTCTGCGCACTCTCTCTGCTTCCTC 63

QY 303 TCAATGGTGTGCCAGCTGTGCCGACACACCTGTACCTGTCTTGGACACACCCAG 362  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 64 TCAAAAGGTGCGTACCAGCTGTGCCGACACACCTGTACCTGTCTTGGACACCCGA 123  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
QY 363 TCCCCACAGGGGGTACCCCTGTGTGGATGGCTGTGGCTGTGTAAAGTGTGTGACGG 422  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 124 TGGCGCTGGGATACCCCTGTGTGGATGGCTGTGGCTGTGTGCGGGTATGTGCACGG 183  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
QY 423 AGGCTGGGGAGTCTCTGGACACACCTGTGTGCGACCCGACGCGCTGGTTTGT 482  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 184 CGGCTGGGGAGCCCTGTGCCCACTTCCAGTCTGTGGACCCGACGCGCTGGTCTCC 243  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
QY 483 CAGCCTGGGGCAGGCCCTGTGGCGCCATGGGGCTGTGTCTTGGATGAGGATGACGCT 542  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 244 CAGCCCGGGCAGACCCGCTGTGGCGGCCCTGTGTCTTGGCAGAGACGACACG 303  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
QY 543 AGCTGTGAGGTGAATGGCGGAGGTACCTGGATGGAGAGACCTTTAAACCCCAATTGACGG 602  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 304 AGCTGTGAGGTGAACGCGCCCTGTATCGGAAGGGAGACCTTCCAGCCCACTGCAGC 363  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
QY 603 GTCCCTGTGCCCTGTGTGATGACGCTGTTCACC 635  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 364 ATCCGCTGCCGCTGTGGAGACGCGGCTTCACC 396  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 14  
AAV29260  
..ID AAV29260 standard; DNA; 1285 BP.  
XX  
XX AC AAV29260;  
XX  
DT 14-SEP-1998 (first entry)  
XX  
DE Human connective tissue growth factor-3 gene.  
XX  
KW Connective tissue growth factor-3; CTGF-3; human; cancer;  
KW arthritis; fibrosis; osteoporosis; diagnosis; therapy; ds.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 9..761 /\*tag= a  
FT sig\_peptide 9..65 /\*tag= b  
FT mat\_peptide 66..758 /\*tag= c  
FT  
XX  
PN WO9821236-A1.  
XX  
XX 22-MAY-1998.  
XX  
XX 08-NOV-1996; 96WO-US17856.  
XX  
XX 08-NOV-1996; 96WO-US17856.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Chopra A, Ebner R, Ruben SM;  
XX WPI; 1998-297864/26.  
DR P-PSDB; AAW37946.  
XX  
XX Novel human connective tissue growth factor 3 gene - useful for the  
PT diagnosis and treatment of e.g. cancer, arthritis, fibrosis,  
PT osteoporosis  
XX  
PS Claim 2; Fig 1A-B; 87pp; English.  
XX  
CC This nucleotide sequence codes for human connective tissue growth  
CC factor-3 (CTGF-3) protein (see AAW37946), a novel member of the  
CC growth factor superfamily. It was discovered in a cDNA library



Fri Aug 1 10:44:14 2003

us-10-010-408-1\_copy\_1\_635.rng

Page 12

Search completed: July 23, 2003, 11:15:17  
Job time : 169.151 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 15:55:06 ; Search time 34.5474 Seconds  
(without alignments)  
5636.882 Million cell updates/sec

Title: US-10-010-408-1\_COPY\_1\_635  
Perfect score: 635  
Sequence: 1 GACGCTTCTGATCTCCAGAG.....GTGATGACGGTGGCTTCACC 635

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_NA.\*

- 1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*
- 2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*
- 3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq.\*
- 4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq.\*
- 5: /cgn2\_6/ptodata/1/ina/PTUS\_COMB.seq.\*
- 6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description        |
|------------|-------|-------------|--------|----|--------------------|
| 1          | 521.4 | 82.1        | 1734   | 4  | US-09-182-145-17   |
| c 2        | 521.4 | 82.1        | 1734   | 4  | Sequence 17, Appli |
| 3          | 273   | 43.0        | 1293   | 4  | US-09-182-145-18   |
| c 4        | 273   | 43.0        | 1293   | 4  | Sequence 18, Appli |
| 5          | 271.4 | 42.7        | 841    | 4  | US-09-182-145-13   |
| 6          | 261.4 | 41.2        | 738    | 4  | Sequence 13, Appli |
| 7          | 87.8  | 13.8        | 1062   | 4  | US-09-182-145-14   |
| 8          | 84    | 13.2        | 1418   | 4  | Sequence 39, Appli |
| 9          | 83.8  | 13.2        | 1480   | 4  | Sequence 38, Appli |
| 10         | 79.2  | 12.5        | 1128   | 2  | US-09-253-316-3    |
| 11         | 78    | 12.3        | 1766   | 4  | US-09-142-569-1    |
| c 12       | 78    | 12.3        | 1766   | 4  | US-08-459-101A-1   |
| 13         | 75.2  | 11.8        | 2830   | 4  | US-09-182-145-9    |
| c 14       | 75.2  | 11.8        | 2830   | 4  | US-09-182-145-10   |
| 15         | 69.8  | 11.0        | 1142   | 4  | US-09-182-145-1    |
| 16         | 69.8  | 11.0        | 1212   | 4  | Sequence 1, Appli  |
| c 17       | 69.8  | 11.0        | 1212   | 4  | Sequence 9, Appli  |
| 18         | 69.8  | 11.0        | 1335   | 4  | Sequence 10, Appli |
| c 19       | 69.8  | 11.0        | 1335   | 4  | Sequence 1, Appli  |
| 20         | 66    | 10.4        | 2075   | 1  | US-09-182-145-2    |
| 21         | 66    | 10.4        | 2075   | 1  | US-09-253-316-1    |
| 22         | 66    | 10.4        | 2075   | 1  | US-08-386-680-1    |
| 23         | 66    | 10.4        | 2075   | 1  | US-08-459-717-1    |
| 24         | 66    | 10.4        | 2075   | 2  | US-08-712-302-1    |
| 25         | 66    | 10.4        | 2075   | 2  | US-08-880-031-1    |
| 26         | 66    | 10.4        | 2075   | 3  | US-09-097-179-1    |
| 27         | 66    | 10.4        | 2075   | 4  | US-09-080-715-1    |
|            |       |             |        |    | Sequence 1, Appli  |
|            |       |             |        |    | Sequence 7, Appli  |

|      |      |      |         |   |                   |                    |
|------|------|------|---------|---|-------------------|--------------------|
| 28   | 66   | 10.4 | 2075    | 5 | PCT-US96-08140-1  | Sequence 1, Appli  |
| 29   | 66   | 10.4 | 2998    | 3 | US-09-054-368-1   | Sequence 1, Appli  |
| 30   | 66   | 10.4 | 2998    | 3 | US-09-054-274-1   | Sequence 1, Appli  |
| 31   | 66   | 10.4 | 2998    | 4 | US-09-056-704-1   | Sequence 1, Appli  |
| 32   | 63.4 | 10.0 | 693     | 4 | US-09-182-145-24  | Sequence 24, Appli |
| 33   | 63.4 | 10.0 | 1101    | 4 | US-09-182-145-29  | Sequence 29, Appli |
| 34   | 63.4 | 10.0 | 1202    | 4 | US-09-182-145-26  | Sequence 23, Appli |
| 35   | 63.4 | 10.0 | 1403    | 4 | US-09-182-145-23  | Sequence 1, Appli  |
| 36   | 62.6 | 9.9  | 2350    | 4 | US-09-187-478-1   | Sequence 1, Appli  |
| 37   | 62.6 | 9.9  | 2350    | 4 | US-09-292-036-1   | Sequence 5, Appli  |
| 38   | 60.8 | 9.6  | 2267    | 4 | US-09-142-569-5   | Sequence 1, Appli  |
| 39   | 55.8 | 8.8  | 4214    | 4 | US-09-122-135-1   | Sequence 1, Appli  |
| 40   | 48.6 | 7.7  | 2541    | 2 | US-08-656-393-1   | Sequence 1, Appli  |
| 41   | 41.4 | 6.5  | 51      | 4 | US-09-182-145-117 | Sequence 117, App  |
| c 42 | 36.6 | 5.8  | 4403765 | 4 | US-09-103-840A-2  | Sequence 1, Appli  |
| c 43 | 36.6 | 5.8  | 4411529 | 4 | US-09-103-840A-1  | Sequence 2, Appli  |
| 44   | 35.2 | 5.5  | 2300    | 3 | US-08-486-343A-2  | Sequence 79, Appli |
| 45   | 34.4 | 5.4  | 87350   | 3 | US-08-781-891-79  |                    |

## ALIGNMENTS

RESULT 1  
US-09-182-145-17  
; Sequence 17, Application US/09182145B  
; Patent No. 6387657  
; GENERAL INFORMATION:  
; APPLICANT: Botstein, David A.  
; APPLICANT: Cohen, Robert  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth J.  
; APPLICANT: Lawrence, David A.  
; APPLICANT: Levine, Arnold J.  
; APPLICANT: Pennica, Diane  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Wood, William I.  
; TITLE OF INVENTION: WSP.POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: P117682  
; CURRENT APPLICATION NUMBER: US/09/182,145B  
; CURRENT FILING DATE: 1998-10-29  
; EARLIER APPLICATION NUMBER: US 60/063,704  
; EARLIER FILING DATE: 1997-10-29  
; EARLIER APPLICATION NUMBER: US 60/073,612  
; EARLIER FILING DATE: 1998-02-04  
; EARLIER APPLICATION NUMBER: US 60/081,695  
; EARLIER FILING DATE: 1998-04-14  
; NUMBER OF SEQ ID NOS: 156  
; SEQ ID NO 17  
; LENGTH: 1734  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-09-182-145-17

|                       |                 |  |            |              |
|-----------------------|-----------------|--|------------|--------------|
| Query Match           | 82.1%           | Score 521.4;   | DB 4;      | Length 1734; |
| Best Local Similarity | 92.0%;          | Pred. No. 1.1e-135;  |            |              |
| Matches 586;          | Conservative 0; | Mismatches 41;   | Indels 10; | Gaps 3;      |
| Qy                    | 3               | CGCTTCTGATCTCCAGAGACCCCTGGGGTGGGACAGGGGCCCTTGGCAAGGCTGCAGCCGC 62 |            |              |
| Db                    | 13              | CGCTCCTGATCTCCAGAGACCCCGGCTGGGACAGGGGCCCTTGGCAGGCTGCAGCTGC 72    |            |              |
| Qy                    | 63              | TG-GGCAGTGGCTGGATGGAGCTCTTATCTACTGGGAAGTGGAGCTAAGAGCTCC 121      |            |              |
| Db                    | 73              | TGTGGCAGTCTGGGATGGAGGCTCTTCTTGTGGGAAGTGGAGGCTGAGAGCTCC 132       |            |              |
| Qy                    | 122             | TGTCAG---CWTGCTAAAGCTTAGCAGCTTGGTGGCTTGGCTTGCACACACTGTCA 178     |            |              |
| Db                    | 133             | TGTCAGGCTCCTGCTCCAAACTCTTGGGACTTGGGCTTGGGCTTGCACACACTGTCA 192    |            |              |
| Qy                    | 179             | GACACCTTCGTGGTGGGCTCCACGCCCTCAGCTTGAAGCTTGCTCCCAAGGG 238         |            |              |

Db 193 GACACCTTCTGGTGGCCTCTCCGGCC-----TCAGGTTTGAAGCTGGCTGCACCAAGG 246

QY 239 ACACGGTGACATGAGGGGAGCCCACTGATCCATCTTCTGGCCACTTCTTCTCTGCTGCT 298

Db 247 ACACGGTGACATGAGGGGAGCCCACTGATCCATCTTCTGGCCACTTCTTCTCTGCTGCT 306

QY 299 TCTCTCAATGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 358

Db 307 TCTCTCAATGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 366

QY 359 CCAGTGGCCACAGGGGTACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 418

Db 367 CCAGTGGCCACAGGGGTACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 426

QY 419 ACAGGAGCTGGGGGAGTCTGCGACACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 478

Db 427 ACAGGAGCTGGGGGAGTCTGCGACACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 486

QY 479 TTGTCAGCTGGGGGAGGCGCTGCGACACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 538

Db 487 TTGTCAGCTGGGGGAGGCGCTGCGACACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 546

QY 539 CGGTAGCTGAGGTGAATGGCGGAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 598

Db 547 CGGTAGCTGAGGTGAATGGCGGAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 606

QY 599 CAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 635

Db 607 CAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 643

RESULT 2

US-09-182-145-18/c

; Sequence 18, Application. US/09182145B

; Patent No. 6387657

; GENERAL INFORMATION:

; APPLICANT: Botstein, David A.

; APPLICANT: Cohen, Robert

; APPLICANT: Goddard, Audrey

; APPLICANT: Gurney, Austin L.

; APPLICANT: Hillan, Kenneth J.

; APPLICANT: Lawrence, David A.

; APPLICANT: Levine, Arnold J.

; APPLICANT: Pennica, Diane

; APPLICANT: Roy, Margaret Ann

; APPLICANT: Wood, William I.

; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME

; FILE REFERENCE: P1176R2

; CURRENT APPLICATION NUMBER: US/09/182,145B

; CURRENT FILING DATE: 1998-10-29

; EARLIER APPLICATION NUMBER: US 60/063,704

; EARLIER FILING DATE: 1997-10-29

; EARLIER APPLICATION NUMBER: US 60/073,612

; EARLIER FILING DATE: 1998-02-04

; EARLIER APPLICATION NUMBER: US 60/081,695

; EARLIER FILING DATE: 1998-04-14

; NUMBER OF SEQ ID NOS: 156

; SEQ ID NO 18

; LENGTH: 1734

; TYPE: DNA

; ORGANISM: Mus musculus

US-09-182-145-18

Query Match 82.1%; Score 521.4; DB 4; Length 1734;

Best Local Similarity 92.0%; Pred. No. 1.1e-135;

Matches 586; Conservative 0; Mismatches 41; Indels 10; Gaps 3;

QY 3 CGCTTCTGATCTCCAGAGACCCCTGGGGTGGACAGGGGCTTGGCAAGGCTGCAGCGCC 62

Db 1722 CGCTCTGATCTCCAGAGACCCCGGGCTGGGACAGGGGCTTGGCGAGGCTGCAGCTGC 1663

QY 63 TG-GCAGTGGCTTGAATGGAGTCTTATTACTGGGAGCTGAGGAGCTAGAGGCTCC 121

Db 1662 TGTGGCAGTAGGTTGGGATGGAGGCTCTTCTTGTGTTGGGAACCTGAGAGAGCTGAGAGGCTCC 1603

QY 122 TGTCAAG--CTTGTCTCTAAAGTCTTAGCACTTGTGTGGCTTGGGCTTTCACACACTGTCA 178

Db 1602 TGTCAAGGCTCTTGTCTTAACTTGTGGCACTTGGGCTTGGGCTTTCACACACTGTCA 1543

QY 179 GACACTTGTGTGGTGGCTTCCACGGCTTACCTTGTGAGTTTGAAGTGTGCTTCCACAGGG 238

Db 1542 GACACTTGTGTGGTGGCTTCCCGGCC-----TCAGGTTTGAAGTGTGCTTCCACAGGG 1489

QY 239 ACACGGTGACATGAGGGGAGCCCACTGATCCATCTTCTGGCCACTTCTTCTCTGCTGCT 298

Db 1488 ACACGGTGACATGAGGGGAGCCCACTGATCCATCTTCTGGCCACTTCTTCTCTGCTGCT 1429

QY 299 TCTCTCAATGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 358

Db 1428 TCTCTCAATGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1369

QY 359 CCAGTGGCCACAGGGGATACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 418

Db 1368 CCAGTGGCCACAGGGGATACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1309

QY 419 ACAGGAGCTGGGGGAGTCTGCGACACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 478

Db 1308 ACAGGAGCTGGGGGAGTCTGCGACACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1249

QY 479 TTGTCAGCTGGGGGAGGCGCTTGGGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 538

Db 1248 TTGTCAGCTGGGGGAGGCGCTTGGGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1189

QY 539 CGGTAGCTGAGGTGAATGGCGGAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 598

Db 1188 CGGGAGCTGTGAGGTGAATGGCGGAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1129

QY 599 CAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 635

Db 1128 CAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1092

RESULT 3

US-09-182-145-13

; Sequence 13, Application. US/09182145B

; Patent No. 6387657

; GENERAL INFORMATION:

; APPLICANT: Botstein, David A.

; APPLICANT: Cohen, Robert

; APPLICANT: Goddard, Audrey

; APPLICANT: Gurney, Austin L.

; APPLICANT: Hillan, Kenneth J.

; APPLICANT: Lawrence, David A.

; APPLICANT: Levine, Arnold J.

; APPLICANT: Pennica, Diane

; APPLICANT: Roy, Margaret Ann

; APPLICANT: Wood, William I.

; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME

; FILE REFERENCE: P1176R2

; CURRENT APPLICATION NUMBER: US/09/182,145B

; CURRENT FILING DATE: 1998-10-29

; EARLIER APPLICATION NUMBER: US 60/063,704

; EARLIER FILING DATE: 1997-10-29

; EARLIER APPLICATION NUMBER: US 60/073,612

; EARLIER FILING DATE: 1998-02-04

; EARLIER APPLICATION NUMBER: US 60/081,695

; EARLIER FILING DATE: 1998-04-14

; NUMBER OF SEQ ID NOS: 156

; SEQ ID NO 13

; LENGTH: 1293

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-182-145-13

Query Match 48.0%; Score 273; DB 4; Length 1293;

Best Local Similarity 80.9%; Pred. No. 1e-66;



|    |      |   |      |
|----|------|---|------|
| Db | 1218 | TCAAGGTGGCTACCCAGCTGTGCCGACACCAATGATCTGCCCTTGGCCACCTCCCCGA  | 1159 |
| Qy | 363  | TGCCACAGGGGTACCCCTGGTCTGGATGGCTGTGGCTGCTAAAGTGTGTGCACGG     | 422  |
| Db | 1158 | TGCCCGCTGGAGTACCCCTGGTCTGGATGGCTGTGGCTGTGCCGGGTATCTGCACGG   | 1099 |
| Qy | 423  | AGGCTGGGGAGTCTCGGACCACTGCATGCTCGCACCCGACGCCAGGCCCTGTTTGT    | 482  |
| Db | 1098 | CGGCTGGGGAGCGCCTCGACCACTCCAGCTCTGGCAGCCGACGAGGCCCTGGTCTGC   | 1039 |
| Qy | 483  | CAGCCTGGGGCAGGCCCTGCGGCCATGGGGCTGTGTCTCTTGGATGAGGATGACGGT   | 542  |
| Db | 1038 | CAGCCCGGGCAGGACCCGCTGGCGGGGGCCCTGTGCTCTTGGCAGAGGACGACAGC    | 979  |
| Qy | 543  | AGCTGTGAGTGAATGCGCGCAGGTACCTGGATGGAGACACCTTAAACCCAATTCGAGG  | 602  |
| Db | 978  | AGCTGTGAGTGAACGCGCCGCTGTATCGGGGAGGGGAGACCTTCACGCCCCACTTCAGC | 919  |
| Qy | 603  | GTCTGTGCGGCTGTGATGACGGTGGCTTCACC                            | 635  |
| Db | 918  | ATCCGCTGCCGCTGCGAGGACGGCGGCTTCACC                           | 886  |

US-09-182-143-39  
; Sequence 39, Application US/09182145B  
; Patent No. 6387657  
; GENERAL INFORMATION:  
; APPLICANT: Borstein, David A.

## RESULT 4

```

: APPLICANT: Hillan, Kenneth J.
: APPLICANT: Lawrence, David A.
: APPLICANT: Levine, Arnold J.
: APPLICANT: Pennica, Diane
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Wood, William I.
: TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
: FILE REFERENCE: P1176R2
: CURRENT APPLICATION NUMBER: US/09/182.145B
: CURRENT FILING DATE: 1998-10-29
: EARLIER APPLICATION NUMBER: US 60/063,704
: EARLIER FILING DATE: 1997-10-29
: EARLIER APPLICATION NUMBER: US 60/073,612
: EARLIER FILING DATE: 1998-02-04
: EARLIER APPLICATION NUMBER: US 60/081,695
: EARLIER FILING DATE: 1998-04-14
: NUMBER OF SEQ ID NOS: 156
: SEQ ID NO 39
: LENGTH: 841
: TYPE: DNA
: ORGANISM: Artificial sequence
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 1-841
: OTHER INFORMATION: Sequence is synthesized.
: Patent NO. 6387657
: US-09-182-145-39

```

|    | Matches | 317;   | Conservative | 0;  | Mismatches | 76; | Indels | 0; | Gaps | 0; |
|----|---------|--|--------------|-----|------------|-----|--------|----|------|----|
| QY | 243     | GGTGACATGAGGGCAGCCCACTGATCCATCTTCTTGGCCACTTCCTCTCCCTCGCTTCTC | TCT          | 302 |            |     |        |    |      |    |
|    |         |  |              |     |            |     |        |    |      |    |
| Db | 6       | GGGCATGAGGCACACCAGAACCCCTCTGGCCCTTCTCCCTCTCTGCCCTCTC         | C            | 65  |            |     |        |    |      |    |
| QY | 303     | TCAATGTGTGTGCCCAAGCTGTGGCGGACACCTGTACCTGTCTGTGGACACACACCCCAG | C            | 362 |            |     |        |    |      |    |
|    |         |  |              |     |            |     |        |    |      |    |
| Db | 66      | TCAAAGTGGGTACCCAGCTGTGCCCGACACCATGTACCTGCCCTGGCCACCTCCOCCA   | A            | 125 |            |     |        |    |      |    |
| QY | 363     | TGCCACAGGGGGTACCCCTGGTCTGGATGGGTGTGGCTGCTGTAAGTGTGTGCACGG    | G            | 422 |            |     |        |    |      |    |

Db 126 TCCCGCTGGAGTACCCCTGGTGGATGGCTGTGGCTGTGGCGGATGTGCACGG 185  
QY 423 AGGCTGGGGAGTCTCGACACCTGCATGCTCGACCCAGCCAGCGGCTGTTGT 482  
Db 186 CGCTGGGGAGCCCTCGACCACTCCACGCTCGGACGCGAGCGGCTGTGTCG 245  
QY 483 CAGCTGGGGAGGCCCTGGGGCCATGGGCTGTGTCTCTTGGATGAGGATGACGGT 542  
Db 246 CAGCCCGGGCAGGACCGGTGGCGGGGCCCTGTGCTCTTGGCAGAGGACGACG 305  
QY 543 AGCTGTAGGTGAATGGCGGAGGTACCTGTGATGAGAGACCTTTAAACCCAAATTCGAGG 602  
Db 306 AGCTGTAGGTGAACGGCGGCTGTATCGGAAGGGAGAGCTTCCAGCCCGCTGCAGC 365  
QY 603 GTCCTGTGCGCTGTGATGACGGTGGCTTACC 635  
Db 366 ATCCGCTCGCGTGGAGGAGCGGCTTACC 398  
RESULT 6  
US-09-182-145-38  
; Sequence 38, Application US/09182145B  
; Patent No. 6387657  
; GENERAL INFORMATION:  
; APPLICANT: Botstein, David A.  
; APPLICANT: Cohen, Robert  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth J.  
; APPLICANT: Lawrence, David A.  
; APPLICANT: Levine, Arnold J.  
; APPLICANT: Pennica, Diane  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Wood, William I.  
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: P1176R2  
; CURRENT APPLICATION NUMBER: US/09/182,145B  
; CURRENT FILING DATE: 1998-10-29  
; EARLIER APPLICATION NUMBER: US 60/063,704  
; EARLIER FILING DATE: 1997-10-29  
; EARLIER APPLICATION NUMBER: US 60/073,612  
; EARLIER FILING DATE: 1998-02-04  
; EARLIER APPLICATION NUMBER: US 60/081,695  
; EARLIER FILING DATE: 1998-04-14  
; NUMBER OF SEQ ID NOS: 156  
; SEQ ID NO 38  
; LENGTH: 738  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-182-145-38

Query Match 41.2%; Score 261.4; DB 4; Length 738;  
Best Local Similarity 81.1%; Pred. No. 1.4e-63;  
Matches 304; Conservative 0; Mismatches 71; Indels 0; Gaps 0;  
QY 261 CCACTGATCCATCTTCTGGCCACTTCTCTCTCTGCTCTCTCAATGGTGTGCCCG 320  
Db 1 CGGAGAGCCACCTCTCTGSCCTTCTCCCTCTCTCTCTCTCAAGGTGGTACCCAG 60  
QY 321 CTGTGCCGAGACCTGTACCTGTGCTGTGGACACACCCAGTGCACAGGGGTACCC 380  
Db 61 CTGTGCCGAGACCATGTGTACCTGTGCCCTTCTCCCTCTCTCTCTCTGAGTACCC 120  
QY 381 CTGGTGTGGATGGCTGTGGTGTGTAAGTGTGTGCACGAGGCTGGGGAGTCCTGC 440  
Db 121 CTGGTGTGGATGGCTGTGGTGTGTCGGGTGTGTGCACCGGCTGGGGAGCCCTTC 180  
QY 441 GACCACATGATGTGTGACACCCAGCCAGGCGCTGGTGTGTGTCAGCCTGGGGAGCCCT 500  
Db 181 GACCAACTCCACGTCTGCGAGCCGACCCAGGCGCTGTCTGCCAGCCCGGGGAGCC 240  
QY 501 GCGGCCATGGGCTGTGTCTCTTGGATGAGGATGACGGTGTGTGAGTGTGATGCC 560

Db 241 GGTGCCGGGGGGCCCTGTGCTCTTGGCAGAGGACGACAGCAGCTGTGAGGTGAACGC 300  
QY 561 CGCAGGTACCTGGATGGAGAGACCTTTAAACCCAAATTCAGGTCCTCTGCGGCTGTGAT 620  
Db 301 CGCTGTATCGGAAGGGAGACCTTCCAGCCCCACTGCAGCATCCGCTGCGGCTGCGAG 360  
QY 621 GACGCTGGCTTACC 635  
Db 361 GACGGCGCTTACC 375  
RESULT 7  
US-09-253-316-3  
; Sequence 3, Application US/09253316  
; Patent No. 639590  
; GENERAL INFORMATION:  
; APPLICANT: Sheppard, Paul O.  
; APPLICANT: Jaspers, Stephen R.  
; TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR HOMOLOGS  
; FILE REFERENCE: 97-75  
; CURRENT APPLICATION NUMBER: US/09/253,316  
; CURRENT FILING DATE: 1999-02-19  
; EARLIER APPLICATION NUMBER: US 60/075,300  
; EARLIER FILING DATE: 1998-02-20  
; NUMBER OF SEQ ID NOS: 34  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 3  
; LENGTH: 1062  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Degenerate sequence  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(1062)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-253-316-3

Query Match 13.8%; Score 87.8; DB 8; Length 1062;  
Best Local Similarity 36.8%; Pred. No. 2.7e-15;  
Matches 114; Conservative 53; Mismatches 143; Indels 0; Gaps 0;  
QY 318 CAGCTGTGCCGACACCTGTACCTGTGCTGTGACACCCAGTGCACAGGGGTA 377  
Db 136 CARTTGTGCAVTGGCCNTGYAARTGYCCNCARCARAARCCNMGTGYCCNCNGNGTN 195  
QY 378 CCCCTGGTGGATGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 437  
Db 196 WSNYTNMTNMGAYGGNTGYAARATHGTGYAARATHGTGYAARATHGTGYAARATHGT 255  
QY 438 TGGACACCTGCTGTCTGCGACCCAGCCAGCGGCTGTGCTGTGCTGTGCTGTGCTGTGCT 497  
Db 256 TGYAARGCNGAYTNTGYAAYCCAYARGGNTYATGYGATYATWSNCTNGAY 315  
QY 498 CTGGCGCCATGGGCTGTGCTGTCTTGGATGAGGATGACGCTGTGCTGTGCTGTGCTGTG 557  
Db 316 MGNCCNMGTAYGARACNGGNTGYGCNTAYTNTGNCNGTNGGNTGYGARTYAY 375  
QY 558 GCGCGCAGGTACCTGGATGGAGAGACCTTTAAACCCAAATTCAGGCTGTGCTGTGCTGTG 617  
Db 376 CARGTNCAYTAYCAAYAGGAGGACGNTTTCARCCNAAYCCNTTNTWSNTGYTNTGY 435  
QY 618 GATGACGGT 627  
Db 436 GTNWSNGNG 445

RESULT 8  
US-09-142-569-3  
; Sequence 3, Application US/09142569  
; Patent No. 641375  
; GENERAL INFORMATION:

| Query Match           | 13.2%           | Score 84;   | DB 4;     | Length 1418; |
|-----------------------|-----------------|---|-----------|--------------|
| Best Local Similarity | 54.5%;          | Pred. No. 3,4e-14;  |           |              |
| Matches 168;          | Conservative 0; | Mismatches 140;   | Indels 0; | Gaps 0;      |
| QY                    | 324             | TGCGGAGACCCCTGTACCTGCTCCTTGGACACACACCCAGTCGCCACAGGGGGGTACCCCTG  | 383       |              |
| Db                    | 199             | TGCGCGCGCTGCCCTGCCACTGCCCCCTGGAGGCGCCCAAGTGCGGCGCGGAGTCGGGCTG   | 258       |              |
| QY                    | 384             | GTCTCTGGATGGCTGTGGCTGCTCTAAAGTGTGTGCACGGAGGCTGGGGGAGTCTCTCGCAC  | 443       |              |
| Db                    | 259             | GTCCGGGACGGCTGCGGCTGCTGTAGGTCGTGGCCNAGCAGCTCAACGAGGACTGTCAGC    | 318       |              |
| QY                    | 444             | CACCTGCATGTCTGCACACCCACGCGGCTGTGTTGTACGCTTGGGGCAGGCCCTTGGC      | 503       |              |
| Db                    | 319             | AAAAACGACGCCCTGCGACCAACACAAAGGGCTGGAATGCAACTTCGSCGCCAGCTCCACC   | 378       |              |
| QY                    | 504             | GGCCATGGGCTGTGTCTCTTTGGATGAGGATGACGGTACGCTGTGAGGTGAATGGCCGC     | 563       |              |
| Db                    | 379             | GCCTCTGAAGGGGATCTGCAGAGCTCAGTCAGAGGGCAGACCCCTGTGGAATAATACTCCAGA | 438       |              |
| QY                    | 564             | AGGTACTTGGATGGAGACGCTTTTAAACCCAAATTTGCAGGGTCTCTGTGCCGCTGTGATGAC | 623       |              |
| Db                    | 439             | ATCTACCAAAACGGGAAAGTTTCCAGCCCAACTGTCAACATCAGTCACATGATTATTGAT    | 498       |              |
| QY                    | 624             | GGTGGCTT  | 631       |              |
| Db                    | 499             | GGCGCGGT  | 506       |              |

|                       |                 |   |           |              |
|-----------------------|-----------------|---|-----------|--------------|
| Query Match           | 13.2%           | Score 83.8;   | DB 4;     | Length 1480; |
| Best Local Similarity | 51.2%;          | Pred. No. 3.9e-14;  |           |              |
| Matches 196;          | Conservative 0; | Mismatches 187;   | Indels 0; | Gaps 0;      |
| Qy                    | 249             | ATGAGGGGACGCCACTGATCCATCTCTTGCCGACTTCCCTTCCTCGCTTCTCTCAATG      | 308       |              |
| Db                    | 180             | ATGAGCTCCAGCACCTTCAGGACGCTGCGTGTGCGCGGTCAACCTTCTCCACTTGACCCAGA  | 239       |              |
| Qy                    | 309             | GTGTGTGCCGAGCTGTCCGGACACCCGTGTACCTGTCTCTTGACACACACCCACGTCGCCA   | 368       |              |
| Db                    | 240             | CTGCGGTCTCCACCTGCCCGCGCGCTGCCACTGCCCTCTGGAGGACCCCAAGTGCCTC      | 299       |              |
| Qy                    | 369             | CAGGGGTACCCCTGGTGTGGATGGTGTGGCTGTCTAAAGTGTGTGCACGGAGGCTG        | 428       |              |
| Db                    | 300             | CCGGAGTCCGGTTGGTCCGGGACGGCTGTGCGGCTGTGTAAAGTGTGCGCTTAACAACATC   | 359       |              |
| Qy                    | 429             | GGGAGTCTCTGCACACCACTGCATGTCTGCGACCCGCCACGGGCTGGTTGTTCAGGCT      | 488       |              |
| Db                    | 360             | AACGAGACTGCACAAACTCAGCCCTGCGACACACACCAAGGGTTGGAATGCAAATTC       | 419       |              |
| Qy                    | 489             | GGGCGAGGCCCTGGCGGCCATGGGGCTGTGTGTCTCTTTGGATGAGGATGACGGTAGCTGT   | 548       |              |
| Db                    | 420             | GGCGCCAGCTCCACCGCTCTGAAGGGATCTGCAGAGCTCAGTCAGAAGGCAGACCCCTGT    | 479       |              |
| Qy                    | 549             | GAGTGTGAATGGCCGACGATACCTGGATGGAGAGACCTTTTAAACCCAAATTCGAGGCTCTGT | 608       |              |

Db 480 GAATATTAAGTCTACCAAAACGGGGAAAGCTTCCACCCCACTGTAAACACACGAG 539

QY 609 TCCCGCTGTGATGACGGTGGCTT 631

Db 540 TGCACATGTATTGATGGCGCGCT 562

## RESULT 10

US-08-459-101A-1

; Sequence 1, Application US/08459101A

; Patent No. 5945300

; GENERAL INFORMATION:

; APPLICANT: LI, ET AL.

; TITLE OF INVENTION: Connective Tissue Growth Factor-2

; NUMBER OF SEQUENCES: 6

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,

; ADDRESSEE: CECCHI, STEWART & OLSTEIN

; STREET: 6 BECKER FARM ROAD

; CITY: ROSELAND

; STATE: NEW JERSEY

; COUNTRY: USA

; ZIP: 07068

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5 INCH DISKETTE

; COMPUTER: IBM PS/2

; OPERATING SYSTEM: MS-DOS

; SOFTWARE: WORD PERFECT 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/459,101A

; FILING DATE: June 2, 1995

; CLASSIFICATION: 536

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/US94/07736

; FILING DATE: 12 JUL 94

; ATTORNEY/AGENT INFORMATION:

; NAME: FERRARO, GREGORY D.

; REGISTRATION NUMBER: 36,134

; REFERENCE/DOCKET NUMBER: 325800-317

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 201-994-1700

; TELEFAX: 201-994-1744

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1128 BASE PAIRS

; TYPE: NUCLEIC ACID

; STRANDEDNESS: SINGLE

; TOPOLOGY: LINEAR

; MOLECULE TYPE: cDNA

US-08-459-101A-1

Query Match 12.5%; Score 79.2; DB 2; Length 1128;

Best Local Similarity 52.1%; Pred. No. 6.7e-13;

Matches 177; Conservative 0; Mismatches 163; Indels 0; Gaps 0;

QY 290 CCTCTGCCTTCTCTCAATGGTGTGGCCGAGCTGTGCCGAGACACCCCTGTACCTGTGCTTGG 349

Db 42 CTTTCCACTTGTACAGGGTGGGGCTCTCCACCTGTGCCCGCTGTACTGCCACTGTGCCCT 101

QY 350 GACACACCCAGTCCCCACAGGGGGTACCCCTGTGCTGTGGATGCTGTGGCTGTGCTGTAA 409

Db 102 GGAGGCGCCCAAGTCCGCGCGGGAGTGGGCTGTGGTGGCGGAGCTGGGCTGGGCTGTGTGTA 161

QY 410 AGTGTGTGCACGGAGGCTGGGGAGTCTCTGCGACACCTGTGATGTCTGCGACCCCGACCA 469

Db 162 GGTCTGCCCAAGCAGCTCAACGAGGACTGCAGAAAAACGACGCCCTCGGACCCACACAA 221

QY 470 GGGCCCTGTGTGTAGCCCTGGGGCAGCCCTGTGGCGGCATGGGGCTGTGTCTCTTGA 529

Db 222 GGGCTGGAATGCAACTTTCGGCGCCAGCTCCACCCCTCTGAAGGGGATCTGCAGAGCTCA 281

QY 530 TGAGGATGACGCTAGCTGTAGGTGAATGGCGGAGGTACCTGGGATGGAGACCTTTAA 589

Db 282 GTACAGAGGGGACACCCCTGTGAATATACTCCAGAATCTACCAAAACGGGAAAGTTTCCA 341

QY 590 ACCCAATTGACGGGTCTGTGCCGTGTGATGACGGTGGC 629

Db 342 GCCCAACTGTAAACATCAGTGCACATGTATTGGATGGCGC 381

## RESULT 11

US-09-182-145-9

; Sequence 9, Application US/09182145B

; Patent No. 6387657

; GENERAL INFORMATION:

; APPLICANT: Botstein, David A.

; APPLICANT: Cohen, Robert

; APPLICANT: Goddard, Audrey

; APPLICANT: Gurney, Austin L.

; APPLICANT: Hillan, Kenneth J.

; APPLICANT: Lawrence, David A.

; APPLICANT: Levine, Arnold J.

; APPLICANT: Pennica, Diane

; APPLICANT: Roy, Margaret Ann

; APPLICANT: Wood, William I.

; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME

; FILE REFERENCE: P1176R2

; CURRENT APPLICATION NUMBER: US/09/182,145B

; CURRENT FILING DATE: 1998-10-29

; EARLIER APPLICATION NUMBER: US 60/063,704

; EARLIER FILING DATE: 1997-10-29

; EARLIER APPLICATION NUMBER: US 60/073,612

; EARLIER FILING DATE: 1998-02-04

; EARLIER APPLICATION NUMBER: US 60/081,695

; EARLIER FILING DATE: 1998-04-14

; NUMBER OF SEQ ID NOS: 156

; SEQ ID NO 9

; LENGTH: 1766

; TYPE: DNA

; ORGANISM: Mus musculus

; FEATURE:

; NAME/KEY: Unsure

; LOCATION: 10

; OTHER INFORMATION: Unknown base

US-09-182-145-9

Query Match 12.3%; Score 78; DB 4; Length 1766;

Best Local Similarity 58.2%; Pred. No. 1.7e-12;

Matches 165; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

QY 316 CCCAGCTGTCCGGACACCCCTGTACCTGTCTCTTGGACACACCACCCAGTCCCGACAGGGGG 375

Db 215 CCGAATTCTGCAAGTGGCCCATGTGAGTCCCCACAATCCCCACCTCGCTGCCACCTGGGGC 274

QY 376 TACCCCTGGTGTGCTGGATGGCTGTGCTGTAAAGTGTGTACAGGAGGCTGGGGGAGT 435

Db 275 TCAGCCTAATACGATGGCTGTGAATGCTGTAAATATGTGCCAGCAGCTTGGGGACA 334

QY 436 CTGCGACACCTGCATGTCTGCGACCCAGCCAGCGGCTGGTGTGTGACGCTGGGGCAG 495

Db 335 ACTGCACAGAGCTGCCCATCTGTGACCCACACCGGGGCTCTACTGCGATTAACAGTGGGG 394

QY 496 GCCCTGGCGGCCATGGGGCTGTGCTCTCTTGGATGAGGATGACCGGTAGCTGTGAGGTGA 555

Db 395 ATCGCCCGAGGTAGCAATAGGAGTGTGTGACACAGTGTGCGGTGTGGGCTGTGCTCTGG 454

QY 556 ATGGCGCGCAGGTACCTTTTAAACCAATTTGACAGGGTCTGTGTGCCGT 615

Db 455 ATGGCGTACGTACCAATGCGAGTCTCTTCCAAACCAACTGCAGAGTACAACCTGTACCT 514

QY 616 GTGATGACGG 625

Db 515 GCATTGATGG 524

## RESULT 12

```

RESULT 13
US-09-142-145-1
; Sequence 1, Application US/09182145B
; Patent No. 6387657
; GENERAL INFORMATION:
; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.

```

RESULT 14  
US-09-182-145-2/c  
; Sequence 2, Application US/09182145B  
; Patent No. 6387657  
; GENERAL INFORMATION:  
; APPLICANT: Botstein, David A.  
; APPLICANT: Cohen, Robert  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth J.  
; APPLICANT: Lawrence, David A.  
; APPLICANT: Levine, Arnold J.  
; APPLICANT: Pennica, Diane  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Wood, William L.  
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: P1176R2  
; CURRENT APPLICATION NUMBER: US/09/182,145B  
; CURRENT FILING DATE: 1998-10-29  
; EARLIER APPLICATION NUMBER: US 60/063,704  
; EARLIER FILING DATE: 1997-10-29

; EARLIER APPLICATION NUMBER: US 60/073,612  
; EARLIER FILING DATE: 1998-02-04  
; EARLIER APPLICATION NUMBER: US 60/081,695  
; EARLIER FILING DATE: 1998-04-14  
; NUMBER OF SEQ ID NOS: 156  
; SEQ ID NO 2  
; LENGTH: 2830  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-182-145-2

Query Match 11.8%; Score 75.2; DB 4; Length 2830;  
Best Local Similarity 52.6%; Pred. No. 1.2e-11;  
Matches 164; Conservative 0; Mismatches 148; Indels 0; Gaps 0;  
QY 316 CCCAGCTGTGCCGGACACCTGTACCTGTCTTGGATGAGGATGGCTGTGATGCTGTAAAAATCTGTGCCAACCAACAGG 264  
DB 2606 CCCAATTTGCGAGTGGCCATGTGAGTGGCCGCCATCCACCCCGCTGCCCGCTGGGGG 2547  
QY 376 TACCCTGTGGTGGTGGCTGTGGCTGTCTGTAAAGTGTGTACAGGAGGCTGGGGGAGT 435  
DB 2546 TCAGCCTCATCACAGATGGCTGTGAGTGTGTAAAGTGTGGCTCAGCAGCTTTGGGACA 2487  
QY 436 CTTGGCAGCACCTGTATCTCGACACCCAGCCAGGGCTGTGTGTAGCTGGGGCAG 495  
DB 2486 ACTGCACGAGGCTGCCATCTGTGACCCGCCACCCGGGCTCTACTGTGACTACAGCGGG 2427  
QY 496 GGCCTGGGGCCATGGGCTGTGTCTTGGATGAGGATGACGCTAGCTGTGAGGTGA 555  
DB 2426 ACCGCCCGAGGTACCATAGAGTGTGTGCACAGTGTGGTGTGGCTGCCGTCTGG 2367  
QY 556 ATGGCCCGCAGGTACTGTGATGAGAGACCTTTAAACCAATTCAGGGTCTCTGTGCCGCT 615  
DB 2366 ATGGGGTGGCTTACAAACAGGCCAGTCTTCCAGCCTAACTGCAAGTACAACTGCACGT 2307  
QY 616 GTGATCAGCGTG 627  
DB 2306 GCATCGACGGCG 2295

RESULT 15  
US-09-253-316-1  
; Sequence 1, Application US/09253316  
; Patent No. 6395890  
; GENERAL INFORMATION:  
; APPLICANT: Sheppard, Paul O.  
; APPLICANT: Jaspers, Stephen R.  
; TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR HOMOLOGS  
; FILE REFERENCE: 97-75  
; CURRENT APPLICATION NUMBER: US/09/253,316  
; CURRENT FILING DATE: 1999-02-19  
; EARLIER APPLICATION NUMBER: US 60/075,300  
; EARLIER FILING DATE: 1998-02-20  
; NUMBER OF SEQ ID NOS: 34  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1  
; LENGTH: 1142  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (17)...(1078)  
US-09-253-316-1

Query Match 11.0%; Score 69.8; DB 4; Length 1142;  
Best Local Similarity 51.1%; Pred. No. 2.7e-10;  
Matches 164; Conservative 0; Mismatches 157; Indels 0; Gaps 0;  
QY 311 GTGTCCCGCAGTGTGCCGGACACCTGTACCTGTCTTGGACACCCACCCAGTGGCCACA 370  
DB 145 GCGTAAACAGTTTGTCTACTGGCCCTGCAATGCCCCCTCAGCAGAGCCCCGTTGCCCTCC 204  
QY 371 GGGGTACCCCTGGTGGTGGTGTGTAAAGTGTGTGACAGGAGGCTGGG 430

DB 205 TGGAGTGGACCTGTGTGAGAGATGGCTGTGGATGCTGTAAAAATCTGTGCCAACCAACAGG 264  
QY 431 GGAGTCCCTGCCAGCACCTGTGATGCTGCGACCCAGCCAGGCGCTGTGTGTGTCAGCCTGG 490  
DB 265 GGAATCTGCAATGAAGCTGACCTCTGTGACCCACACAAAGGCTGTATTTGTGACTACTC 324  
QY 491 GGCAGGCCCTGGGGCCATGGGGCTGTGTCTTGGATGAGGATGACGGTGTGCTGA 550  
DB 325 AGTAGACAGGCTAGGTACGAGACTGGAGTGTGTCATACCTTTGTAGCTGTGGGTGGCA 384  
QY 551 GGTGAATGGCCGAGGTACCTGTGGATGGAGAGACCTTTAAACCAATTCAGGGTCTGTG 610  
DB 385 GTTCAACCAAGTACATTATCATTAATGGCCAAGTGTTCAGGCCAACCCCTTTGTCAGCTG 444  
QY 611 CCGCTGTGATGAGGCTGTGCTT 631  
DB 445 CCTCTGTGTGAGTGGGGCCAT 465

Search completed: July 28, 2003, 21:35:57  
Job time : 37.7474 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 15:36:41 ; Search time 156.533 Seconds  
(without alignments)  
8368.853 Million cell updates/sec

Title: US-10-010-408-1\_copy\_1\_635

Perfect score: 635

Sequence: 1 GACGCTTCTGATCCAGAG.....GTGATGACGGTGCTTCACC 635

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1439767 seqs, 1031500376 residues

Total number of hits satisfying chosen parameters: 2879534

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published\_Applications\_NA:\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
10: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq1:\*  
11: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq2:\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq3:\*  
13: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
15: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*  
16: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*  
17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description       |
|------------|-------|-------------|--------|----|-------------------|
| 1          | 635   | 100.0       | 1708   | 15 | US-10-010-408-1   |
| 2          | 521.4 | 82.1        | 1734   | 15 | US-10-112-267-17  |
| 3          | 521.4 | 82.1        | 1734   | 15 | US-10-112-267-18  |
| 4          | 387   | 60.9        | 753    | 15 | US-10-010-408-3   |
| 5          | 318   | 50.1        | 681    | 15 | US-10-010-408-12  |
| 6          | 278.4 | 43.8        | 1337   | 11 | US-09-915-582-30  |
| 7          | 278.2 | 43.8        | 1352   | 11 | US-09-915-582-14  |
| 8          | 273   | 43.0        | 1266   | 14 | US-10-137-866-319 |
| 9          | 273   | 43.0        | 1266   | 14 | US-10-146-726-319 |
| 10         | 273   | 43.0        | 1266   | 14 | US-10-146-727-319 |
| 11         | 273   | 43.0        | 1266   | 14 | US-10-146-788-319 |
| 12         | 273   | 43.0        | 1266   | 14 | US-10-152-380-319 |
| 13         | 273   | 43.0        | 1266   | 14 | US-10-153-934-319 |
| 14         | 273   | 43.0        | 1266   | 15 | US-10-028-072-319 |
| 15         | 273   | 43.0        | 1266   | 15 | US-10-121-049-319 |
| 16         | 273   | 43.0        | 1266   | 15 | US-10-123-904-319 |
| 17         | 273   | 43.0        | 1266   | 15 | US-10-123-904-319 |

|    |     |      |      |    |                    |                   |
|----|-----|------|------|----|--------------------|-------------------|
| 17 | 273 | 43.0 | 1266 | 15 | US-10-140-470-319  | Sequence 319, App |
| 18 | 273 | 43.0 | 1266 | 15 | US-10-175-746-319  | Sequence 319, App |
| 19 | 273 | 43.0 | 1266 | 15 | US-10-176-918-319  | Sequence 319, App |
| 20 | 273 | 43.0 | 1266 | 15 | US-10-176-921-319  | Sequence 319, App |
| 21 | 273 | 43.0 | 1266 | 15 | US-10-137-865-319  | Sequence 319, App |
| 22 | 273 | 43.0 | 1266 | 15 | US-10-140-474-319  | Sequence 319, App |
| 23 | 273 | 43.0 | 1266 | 15 | US-10-142-431-319  | Sequence 319, App |
| 24 | 273 | 43.0 | 1266 | 15 | US-10-143-114-319  | Sequence 319, App |
| 25 | 273 | 43.0 | 1266 | 15 | US-10-140-002-319  | Sequence 319, App |
| 26 | 273 | 43.0 | 1266 | 15 | US-10-142-419-319  | Sequence 319, App |
| 27 | 273 | 43.0 | 1266 | 15 | US-10-123-262-319  | Sequence 319, App |
| 28 | 273 | 43.0 | 1266 | 15 | US-10-142-423-319  | Sequence 319, App |
| 29 | 273 | 43.0 | 1266 | 15 | US-10-121-050-319  | Sequence 319, App |
| 30 | 273 | 43.0 | 1266 | 15 | US-10-141-755-319  | Sequence 319, App |
| 31 | 273 | 43.0 | 1266 | 15 | US-10-143-032-319  | Sequence 319, App |
| 32 | 273 | 43.0 | 1266 | 15 | US-10-123-108-319  | Sequence 319, App |
| 33 | 273 | 43.0 | 1266 | 15 | US-10-123-236-319  | Sequence 319, App |
| 34 | 273 | 43.0 | 1266 | 15 | US-10-123-261-319  | Sequence 319, App |
| 35 | 273 | 43.0 | 1266 | 15 | US-10-140-921-319  | Sequence 319, App |
| 36 | 273 | 43.0 | 1266 | 15 | US-10-140-928-319  | Sequence 319, App |
| 37 | 273 | 43.0 | 1266 | 15 | US-10-121-045-319  | Sequence 319, App |
| 38 | 273 | 43.0 | 1266 | 15 | US-10-123-292-319  | Sequence 319, App |
| 39 | 273 | 43.0 | 1266 | 15 | US-10-123-903-319  | Sequence 319, App |
| 40 | 273 | 43.0 | 1266 | 15 | US-10-124-819-319  | Sequence 319, App |
| 41 | 273 | 43.0 | 1266 | 15 | US-10-124-822-319  | Sequence 319, App |
| 42 | 273 | 43.0 | 1266 | 15 | US-10-140-925-319  | Sequence 319, App |
| 43 | 273 | 43.0 | 1266 | 15 | US-10-160-498-319  | Sequence 319, App |
| 44 | 273 | 43.0 | 1266 | 15 | US-10-124-824-319  | Sequence 319, App |
| 45 | 273 | 43.0 | 1266 | 15 | US-10-127-825A-319 | Sequence 319, App |

#### ALIGNMENTS

#### RESULT 1

US-10-010-408-1  
; Sequence 1, Application US/10010408  
; Publication No. US20020165185A1  
; GENERAL INFORMATION:

APPLICANT: John J. Castellet, Jr.  
TITLE OF INVENTION: No. US20020165185A1e1 Heparin-Induced CCN-Like Molecules and Uses Therefor  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD, LLP  
STREET: 28 State Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/010,408

FILING DATE: 07-Dec-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/044,273

FILING DATE: March 19, 1998

APPLICATION NUMBER: <Unknown>

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Amy E. Mandragouras

REGISTRATION NUMBER: 36,207

REFERENCE/DOCKET NUMBER: MBI-004

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617)227-7400

TELEFAX: (617)742-4214

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

; LENGTH: 1708 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 249..1001  
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-10-010-408-1

Query Match 100.0%; Score 635; DB 15; Length 1708;  
Best Local Similarity 100.0%; Pred. No. 4.8e-173;  
Matches 635; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGCTTCATCTCCAGAGGACCTGGGGTGGGACAGGGGCTTGGCAAGGCTGCAGCC 60  
DB 1 GACGCTTCATCTCCAGAGGACCTGGGGTGGGACAGGGGCTTGGCAAGGCTGCAGCC 60

QY 61 GCTGGCAGTGGCTTGGGAATGAGGCTCTTTATTACTGGGAACCTGAGGAGCTAAGAGGCTC 120  
DB 61 GCTGGCAGTGGCTTGGGAATGAGGCTCTTTATTACTGGGAACCTGAGGAGCTAAGAGGCTC 120

QY 121 CTGTACGCTTGTCTTAAGTCTTAGCACTTGTGGTGGCTTGGGCTTCACACTGTCA 180  
DB 121 CTGTACGCTTGTCTTAAGTCTTAGCACTTGTGGTGGCTTGGGCTTCACACTGTCA 180

QY 181 CACCTTCGTGGTGGCTCCACGGCTCACCTTCAGGTTTGAAGCTGGCTCCACAAGGAC 240  
DB 181 CACCTTCGTGGTGGCTCCACGGCTCACCTTCAGGTTTGAAGCTGGCTCCACAAGGAC 240

QY 241 ACGGTGACATGAGGGGACGCCACCTGATCCATCTTCTGGCCACTTCTTCCCTGCCCTTC 300  
DB 241 ACGGTGACATGAGGGGACGCCACCTGATCCATCTTCTGGCCACTTCTTCCCTGCCCTTC 300

QY 301 TCTCAATGGTGTGCCAGCTGTGCCGACACCTGTACCTGTCTTGGACACCAACCC 360  
DB 301 TCTCAATGGTGTGCCAGCTGTGCCGACACCTGTACCTGTCTTGGACACCAACCC 360

QY 361 AGTGCCACAGGGGTACCCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 420  
DB 361 AGTGCCACAGGGGTACCCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 420

QY 421 GGAGGCTGGGAGTCTCGACACCTGATGATGATGATGATGATGATGATGATGATGATGAT 480  
DB 421 GGAGGCTGGGAGTCTCGACACCTGATGATGATGATGATGATGATGATGATGATGATGAT 480

QY 481 GTACGCTGGGACGCGCTGGCGCCATGGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 540  
DB 481 GTACGCTGGGACGCGCTGGCGCCATGGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 540

QY 541 GTAGCTGTGAGTGAATGGCCAGGTACCTGGATGGAGAGACCTTTAAACCCAAATTGCA 600  
DB 541 GTAGCTGTGAGTGAATGGCCAGGTACCTGGATGGAGAGACCTTTAAACCCAAATTGCA 600

QY 601 GGGTCTGTGCGCTGTGATGACGCTGGCTTCAAC 635  
DB 601 GGGTCTGTGCGCTGTGATGACGCTGGCTTCAAC 635

## RESULT 2

US-10-112-267-17  
; Sequence 17, Application US/10112267  
; Publication No. US2003006878A1  
; GENERAL INFORMATION:  
; APPLICANT: Botstein, David A.  
; APPLICANT: Cohen, Robert  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth J.  
; APPLICANT: Lawrence, David A.  
; APPLICANT: Levine, Arnold J.  
; APPLICANT: Pennica, Diane

; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Wood, William I.  
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: P1176R2  
; CURRENT APPLICATION NUMBER: US/10/112,267  
; CURRENT FILING DATE: 2002-03-27  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/182,145B  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-29  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/063,704  
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-10-29  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/073,612  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-04  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/081,695  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-14  
; NUMBER OF SEQ ID NOS: 156  
; SEQ ID NO 17  
; LENGTH: 1734  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; US-10-112-267-17

Query Match 82.1%; Score 521.4; DB 15; Length 1734;  
Best Local Similarity 92.0%; Pred. No. 2.3e-140;  
Matches 586; Conservative 0; Mismatches 41; Indels 10; Gaps 3;

QY 3 CGCTTCTGATCTCCAGAGGACCTTGGGGTGGGACAGGGGCTTGGCAAGGCTGCAGCCGC 62  
DB 13 CGCTTCTGATCTCCAGAGGACCTTGGGGTGGGACAGGGGCTTGGCGAGGCTGCAGCTGC 72

QY 63 TG-GCAGTGGCTTGGATGGAGGCTTTATTACTGGGAACCTGAGAGGCTAAGAGGCTCC 121  
DB 73 TGTGCAGTAGCTGGGATGGAGGCTTTTCTTGTGGGAACCTGAGAGGCTGAGAGGCTCC 132

QY 122 TGTGAG---CTGTGCTTAAAGCTTAGCACTTGTGGTGGTGGTGGTGGTGGTGGTGGTGG 178  
DB 133 TGTGAGGCTTCTTAAAGCTTAGCACTTGTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTT 192

QY 179 GACACCTTGTGGTGGCTTCCACGGCTCACCTTCAGGTTTGAAGCTTGGCTTCCCAAGGG 238  
DB 193 GACACCTTGTGGTGGCTTCCCTCGGCC-----TCAGGTTTGAAGCTTCCCAAGGG 246

QY 239 ACACGGTGCATGAGGGGACGCCCACTGATCATCTTCTGGCCACTTCTTCTCTGCT 298  
DB 247 ACACGGTGCATGAGGGGACGCCCACTGATCATCTTCTGGCCACTTCTTCTCTGCT 306

QY 299 TCTCTCAATGGTGGTGGCTTCCAGCTGTGCCGACACCTGTACCTGTCTTGGACACCA 358  
DB 307 TCTCTCAATGGTGGTGGCTTCCAGCTGTGCCGACACCTGTGCTGTCTTGGACACCA 366

QY 359 CCAGTGGCCACAGGGGTTACCCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 418  
DB 367 CCAGTGGCCACAGGGGTTACCCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 426

QY 419 ACGGAGGCTGGGGAGTCTCTGCGACACCTGATCTGCGACCCACCGCCAGCGGCTGGT 478  
DB 427 ACGGAGGCTGGGGAGTCTCTGCGACACCTGATCTGCGACCCACCGCCAGCGGCTGGT 486

QY 479 TTGTGAGGCTGGGGAGGCTTGGGGGCTGGGGGCTGGGGGCTGGGGGCTGGGGGCTGGG 538  
DB 487 TTGTGAGGCTGGGGAGGCTTGGGGGCTGGGGGCTGGGGGCTGGGGGCTGGGGGCTGGG 546

QY 539 CGGTAGCTGTGAGGTGAATGGCCGAGGTACCTGATGGAGAGACCTTTAAACCCAAATTG 598  
DB 547 CGGTAGCTGTGAGGTGAATGGCCGAGGTACCTGATGGAGAGACCTTTAAACCCAAATTG 606

QY 599 CAGGCTCTGTGCGCTGTGATGACGCTGGCTTCAAC 635  
DB 607 CAGGCTCTGTGCGCTGTGATGACGCTGGCTTCAAC 643

## RESULT 3

US-10-112-267-18/c  
; Sequence 18, Application US/10112267



Publication No. US2003006878A1  
GENERAL INFORMATION:  
APPLICANT: Botstein, David A.  
APPLICANT: Cohen, Robert  
APPLICANT: Goddard, Audrey  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: Lawrence, David A.  
APPLICANT: Levine, Arnold J.  
APPLICANT: Pennica, Diane  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME  
FILE REFERENCE: P1176R2  
CURRENT FILING DATE: 2002-03-27  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/182,145B  
PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-29  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/063,704  
PRIOR FILING DATE: EARLIER FILING DATE: 1997-10-29  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/073,612  
PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-04  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/081,695  
PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-14  
NUMBER OF SEQ ID NOS: 156  
SEQ ID NO 18  
LENGTH: 1734  
TYPE: DNA  
ORGANISM: Mus musculus  
US-10-112-267-18

Query Match 82.1%; Score 521.4; DB 15; Length 1734;  
Best Local Similarity 92.0%; Pred. No. 2.3e-140;  
Matches 586; Conservative 0; Mismatches 41; Indels 10; Gaps 3;  
QY 3 CGCTCTGATCTCCAGAGACCCCTGGGTGGGACAGGGCCCTTGCAAGGTGCGAGCGC 62  
DB 1722 CGCTCTGATCTCCAGAGACCCCGGGCTGGGACAGGGCCCTTGCGAGGTGCGACGTG 1663  
QY 63 TG-GCAGGTGGCTGGAATGGAGGTCTTTATTACTTGGAACTGAGAGCTAAGAGGCTCC 121  
DB 1662 TGTGCAAGTGGTGGATGGAGGTCTTCTTGTGGAACGTGAGAGCTGAGAGGCTCC 1603  
QY 122 TGTCAG-CTTGCTCTAAAGTCTTAGCACTTGTGGTGGCTTGGCTTTCACACACTGTCA 178  
DB 1602 TGTCAAGGCTCCTGTCTAAACTCTTGGCACTTGGCGTGGCTTGGCTTTCACACACTGTCA 1543  
QY 179 GACACCTTCGGTGGCTCCAGGCGCTCACCTTCAGGTGTTGAAGCTGGCTCCACAAGGG 238  
DB 1542 GACACCTTCTTGGTGGCTCTCTCGGCC-----TCAGGTTTGAAGCTGGCTCCACAAGGG 1489  
QY 239 ACACGGTGACATGAGGGGCGAGCCCACTGATCCATCTTCTGGCCACTTCTCTCTGCT 298  
DB 1498 ACACGGTGACATGAGGGGCGAGCCCACTGATCCATCTTCTGGCCATCTCTCTCTGCT 1429  
QY 299 TCTCTCAATGTTGTGCTCCAGCTGTGCGGACACCCCTGTACCTGTCTTGGACACCACC 358  
DB 1428 TCTCTCAATGTTGTATTCCAGCTGTGCGGACACCCCTGTGCTGTCTTGGACACCACC 1369  
QY 359 CAAGTCCCAACAGGGGGTACCCCTGGTGTGATGGCTGTGCTGCTGTAAAGTGTGTGC 418  
DB 1368 CAAGTCCCAACAGGGGGTACCCCTGGTGTGATGGCTGTGCTGTGCTGTGCTGTGCTG 1309  
QY 419 AGGAGGCTGGGGAGTCTTGGGACCACTGCATCTCTGCGACCCAGCCAGGCGCTGGT 478  
DB 1308 AGGAGGCTGGGGAGTCTTGGGACCACTGCATCTCTGCGACCCAGCCAGGCGCTGGT 1249  
QY 479 TTGTGAGCTGGGGAGGCTTGGGCGCATGGGCTGTGTCTCTCTTGGATGAGGATGA 538  
DB 1248 TTGTGAGCTGGGGAGGCTTGGGCGCATGGGCTGTGTCTCTCTTTCGAGAGAGATGA 1189  
QY 539 CGGTAGCTGTGAGGTGAATGGCGCGAGGTACCTTGGATGGAGAGACCTTTAAACCCCAATTG 598  
DB 1189 CGGTAGCTGTGAGGTGAATGGCGCGAGGTACCTTGGATGGAGAGACCTTTAAACCCCAATTG 598

Db 1188 CGGAGCTGTGAGGTGAATGGCCGAGCTACCTGGATGGGAGACCTTTAAACCCCAATTG 1129  
QY 599 CAGGTCCTGTGCGGCTGTGATGACGGTGGCTTACCC 635  
DB 1128 CAGGTTTGTGCGGCTGTGATGACGGTGGCTTACCC 1092  
RESULT 4  
US-10-010-408-3  
Sequence 3, Application US/10010408  
Publication No. US20020165185A1  
GENERAL INFORMATION:  
APPLICANT: John J. Castellot, Jr.  
TITLE OF INVENTION: No. US20020165185A1 Heparin-Induced CCN-Like Molecules and Uses Therefor  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD, LLP  
STREET: 28 State Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/010,408  
FILING DATE: 07-Dec-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/044,273  
FILING DATE: March 19, 1998  
APPLICATION NUMBER: <Unknown>  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Amy E. Mandragouras  
REGISTRATION NUMBER: 36,207  
REFERENCE/DOCKET NUMBER: MBI-004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)742-4214  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 753 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..750  
SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
US-10-010-408-3  
Query Match 60.9%; Score 387; DB 15; Length 753;  
Best Local Similarity 100.0%; Pred. No. 1e-101;  
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 249 ATGAGGGGCGAGCCCACTGATCCATCTTCTGGCCACTTCTTCTGCTCTCTCTCTCAATG 308  
DB 1 ATGAGGGGCGAGCCCACTGATCCATCTTCTGGCCACTTCTTCTGCTCTCTCTCTCAATG 60  
QY 309 GTGTGTGCCCCAGCTGTGCGGACACCCCTGTACCTGTCTCTTGGACACCACCCAGTGCCCA 368  
DB 61 GTGTGTGCCCCAGCTGTGCGGACACCCCTGTACCTGTCTCTTGGACACCACCCAGTGCCCA 120  
QY 369 CAGGGGTACCCCTGGTGGTGGCTGTGAGGTGTGCTGTAAAGTGTGTGACGAGGCTG 428  
DB 121 CAGGGGTACCCCTGGTGGTGGCTGTGAGGTGTGCTGTAAAGTGTGTGACGAGGCTG 180

|  |     |   |     |
|--|-----|---|-----|
| Db   | 1   | CAGCTGTGCCGGACACCCCTGTACCTGTCTTTGGACACCCACCCAGTGCACACAGGGGGTA   | 60  |
| QY   | 378 | CCCTCGGTGCTGGATGGCTGTGGCTGTGTTAAAGTGTGTGCACGGAGGCTGGGGGAGTCC    | 437 |
| Db   | 61  | CCCTTGTGTGGATGGCTGTGGCTGTGTTAAAGTGTGTGCACGGAGGCTGGGGGAGTCC      | 120 |
| QY   | 438 | TCCGACCACTGCTGATGTTGCGACCCAGCCAGGCGCTGTTTGTTCAGCCTGGGGCAGGC     | 497 |
| Db   | 121 | TGCGACCACTGCTGATGTTGCGACCCAGCCAGGCGCTGTTTGTTCAGCCTGGGGCAGGC     | 180 |
| QY   | 498 | CCTGGCGGCATGGGGCTGTGTCTCTTGTGATGAGATGACGGTAGCTGTGAGGTGAAT       | 557 |
| Db   | 181 | CCTGGCGGCATGGGGCTGTGTCTCTTGTGATGAGATGACGGTAGCTGTGAGGTGAAT       | 240 |
| QY   | 558 | GGCCCGAGGTACGCTGGATGGAGAGACCTTTTAAACCCAAATTCAGGGTCTCTGTCCCGCTGT | 617 |
| Db   | 241 | GGCCCGAGGTACGCTGGATGGAGAGACCTTTTAAACCCAAATTCAGGGTCTCTGTCCCGCTGT | 300 |
| QY   | 618 | GATGACGGTGGCTTCACC  | 635 |
| Db   | 301 | GATGACGGTGGCTTCACC  | 318 |
| RESULT 6   |     |   |     |
| US-09-915-582-30   |     |   |     |
| ; Sequence 30, Application US/09915582                       |     |   |     |
| ; Patent No. US20020120103A1                                 |     |   |     |
| GENERAL INFORMATION:   |     |   |     |
| ; APPLICANT: Rosen et al.                                    |     |   |     |
| ; TITLE OF INVENTION: 17 Human Secreted Proteins             |     |   |     |
| ; FILE REFERENCE: PS723P1                                    |     |   |     |
| ; CURRENT APPLICATION NUMBER: US/09/915,582                  |     |   |     |
| ; CURRENT FILING DATE: 2001-07-27                            |     |   |     |
| ; PRIOR APPLICATION NUMBER: PCT/US01/01431                   |     |   |     |
| ; PRIOR FILING DATE: 2001-01-17                              |     |   |     |
| ; PRIOR APPLICATION NUMBER: 60/179,065                       |     |   |     |
| ; PRIOR FILING DATE: 2000-01-31                              |     |   |     |
| ; PRIOR APPLICATION NUMBER: 60/180,628                       |     |   |     |
| ; PRIOR FILING DATE: 2000-02-04                              |     |   |     |
| ; PRIOR APPLICATION NUMBER: 60/231,968                       |     |   |     |
| ; PRIOR FILING DATE: 2000-09-12                              |     |   |     |
| ; NUMBER OF SEQ ID NOS: 97                                   |     |   |     |
| ; SOFTWARE: PatentIn Ver. 2.0                                |     |   |     |
| ; SEQ ID NO 30   |     |   |     |
| ; LENGTH: 1337   |     |   |     |
| ; TYPE: DNA  |     |   |     |
| ; ORGANISM: Homo sapiens.                                    |     |   |     |
| ; FEATURE:   |     |   |     |
| ; NAME/KEY: SITE.  |     |   |     |
| ; LOCATION: (1337)   |     |   |     |
| ; OTHER INFORMATION: n equals a,t,g, or c                    |     |   |     |
| US-09-915-582-30   |     |   |     |
| Query Match 43.8%; Score 278.4; DB 11; Length 1337;          |     |   |     |
| Best Local Similarity 78.4%; Pred. No. 1.9e-70;              |     |   |     |
| Matches 345; Conservative 3; Mismatches 84; Indels 8; Gaps 1 |     |   |     |
| QY   | 196 | CTTCACGGCTCACTTCAGGTTTGAAGCTGGCTTCACAAAGGACACAGGTGACATGAGG      | 255 |
| Db   | 2   | CTTCACAGTTTCACCTTCAGGCTCAAAARCTGGSTCTGCA-----GGGACATGAGAG       | 53  |
| QY   | 256 | GCAGCCCACTGATCATCTCTCGCCACTTCCTTCCTCTGCTTCCTCAATGGTGTGTG        | 315 |
| Db   | 54  | GCACACCGAAGACCCACCTCTCGCCCTTCCTTCCTTCCTTCCTCAAGGTGCGTA          | 113 |
| QY   | 316 | CCAGCTGTGGCGGACACCTGTACCTGTCTTTGGACACACCCACCTGACAGGGG           | 375 |
| Db   | 114 | CCAGCTGTGGCGGACACCATGTACCTGCCCTGGCCACCTCCCGATGCCCGTGGAG         | 173 |
| QY   | 376 | TACCCCTGGTGTGATGGCTGTGGCTGTGTAAGTGTGTGCACGAGGCTGGGGGAGT         | 435 |
| Db   | 174 | TACCCCTGGTGTGATGGCTGTGGCTGTGTAAGTGTGTGCACGAGGCTGGGGGAGC         | 233 |



;; PRIOR FILING DATE: 1997-10-27  
;; PRIOR APPLICATION NUMBER: 60/063329  
;; PRIOR FILING DATE: 1997-10-27  
;; PRIOR APPLICATION NUMBER: 60/063550  
;; PRIOR FILING DATE: 1997-10-28  
;; PRIOR APPLICATION NUMBER: 60/063561  
;; PRIOR FILING DATE: 1997-10-28  
;; PRIOR APPLICATION NUMBER: 60/063704  
;; PRIOR FILING DATE: 1997-10-29  
;; PRIOR APPLICATION NUMBER: 60/063733  
;; PRIOR FILING DATE: 1997-10-29  
;; PRIOR APPLICATION NUMBER: 60/063735  
;; PRIOR FILING DATE: 1997-10-29  
;; PRIOR APPLICATION NUMBER: 60/063738  
;; PRIOR FILING DATE: 1997-10-29  
;; PRIOR APPLICATION NUMBER: 60/063755  
;; PRIOR FILING DATE: 1997-10-17  
;; PRIOR APPLICATION NUMBER: 60/064248  
;; PRIOR FILING DATE: 1997-11-03  
;; PRIOR APPLICATION NUMBER: 60/064809  
;; PRIOR FILING DATE: 1997-11-07  
;; PRIOR APPLICATION NUMBER: 60/065186  
;; PRIOR FILING DATE: 1997-11-12  
;; PRIOR APPLICATION NUMBER: 60/065846  
;; PRIOR FILING DATE: 1997-11-17  
;; PRIOR APPLICATION NUMBER: 60/066364  
;; PRIOR FILING DATE: 1997-11-21  
;; PRIOR APPLICATION NUMBER: 60/066453  
;; PRIOR FILING DATE: 1997-11-24  
;; PRIOR APPLICATION NUMBER: 60/066511  
;; PRIOR FILING DATE: 1997-11-24  
;; PRIOR APPLICATION NUMBER: 60/066770  
;; PRIOR FILING DATE: 1997-11-24  
;; PRIOR APPLICATION NUMBER: 60/069212  
;; PRIOR FILING DATE: 1997-12-11  
;; PRIOR APPLICATION NUMBER: 60/069278  
;; PRIOR FILING DATE: 1997-12-11  
;; PRIOR APPLICATION NUMBER: 60/069334  
;; PRIOR FILING DATE: 1997-12-11  
;; PRIOR APPLICATION NUMBER: 60/069694  
;; PRIOR FILING DATE: 1997-12-16  
;; PRIOR APPLICATION NUMBER: 60/072320  
;; PRIOR FILING DATE: 1998-01-23  
;; PRIOR APPLICATION NUMBER: 60/073612  
;; PRIOR FILING DATE: 1998-02-04  
;; PRIOR APPLICATION NUMBER: 60/074086  
;; PRIOR FILING DATE: 1998-02-09  
;; PRIOR APPLICATION NUMBER: 60/074092  
;; PRIOR FILING DATE: 1998-02-09  
;; PRIOR APPLICATION NUMBER: 60/077791  
;; PRIOR FILING DATE: 1998-03-12  
;; PRIOR APPLICATION NUMBER: 60/078910  
;; PRIOR FILING DATE: 1998-03-20  
;; PRIOR APPLICATION NUMBER: 60/079294  
;; PRIOR FILING DATE: 1998-03-25  
;; PRIOR APPLICATION NUMBER: 60/079663  
;; PRIOR FILING DATE: 1998-02-27  
;; PRIOR APPLICATION NUMBER: 60/079728  
;; PRIOR FILING DATE: 1998-03-27  
;; PRIOR APPLICATION NUMBER: 60/080165  
;; PRIOR FILING DATE: 1998-03-31  
;; PRIOR APPLICATION NUMBER: 60/081203  
;; PRIOR FILING DATE: 1998-04-09  
;; PRIOR APPLICATION NUMBER: 60/081229  
;; PRIOR FILING DATE: 1998-04-09  
;; PRIOR APPLICATION NUMBER: 60/081695  
;; PRIOR FILING DATE: 1998-04-14  
;; PRIOR APPLICATION NUMBER: 60/081817  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/081818  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/082999  
;; PRIOR FILING DATE: 1998-04-24

;; PRIOR APPLICATION NUMBER: 60/083322  
;; PRIOR FILING DATE: 1998-04-28  
;; PRIOR APPLICATION NUMBER: 60/083545  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/084600  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084627  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084637  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/085149  
;; PRIOR FILING DATE: 1998-05-12  
;; PRIOR APPLICATION NUMBER: 60/085323  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085338  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085339  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085579  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085697  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085704  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/086414  
;; PRIOR FILING DATE: 1998-05-22  
;; PRIOR APPLICATION NUMBER: 60/086430  
;; PRIOR FILING DATE: 1998-05-22  
;; PRIOR APPLICATION NUMBER: 60/087106  
;; PRIOR FILING DATE: 1998-05-28  
;; PRIOR APPLICATION NUMBER: 60/088026  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088730  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088741  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088810  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088858  
;; PRIOR FILING DATE: 19/98-06-11  
;; PRIOR APPLICATION NUMBER: 60/089532  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089599  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089907  
;; PRIOR FILING DATE: 1998-06-18  
;; PRIOR APPLICATION NUMBER: 60/089947  
;; PRIOR FILING DATE: 1998-06-19  
;; PRIOR APPLICATION NUMBER: 60/090349  
;; PRIOR FILING DATE: 1998-06-23  
;; PRIOR APPLICATION NUMBER: 60/090429  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090445  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090538  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090863  
;; PRIOR FILING DATE: 1998-06-26  
;; PRIOR APPLICATION NUMBER: 60/091360  
;; PRIOR FILING DATE: 1998-07-01  
;; PRIOR APPLICATION NUMBER: 60/091519  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091982

Query Match 43.0%; Score 273; DB 14; Length 1266;

Best Local Similarity 80.9%; Pred. No. 6.8e-69;

Matches 318; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 243 GGTGACATGAGGGGAGGCCACTGATCCATCTTCTGGGCACATTCCTCTGCTGCTCTC 302

DB 4 GGGGACATGAGAGGACACACCCGAGACCCACCTCTGCTGCTCTCCTCTGCTGCTCTC 63

QY 303 TCAATGGTGTGCGCAGCTGTGCGGAGACCCCTGTACTGTCTTGGACACACCCCGAG 362

Db 64 TCAAGGTGGCTACCCAGCTGTGCCGACACCATGTACCTGGCCCTGCCACCTCCCCGA 123  
Qy 363 TGCCACAGAGGGTACCCCTGGTCTGATGGCTGTGGCTGCTGCTAAAGTGTGTCACGG 422  
Db 124 TGCCCGCTGGAGTACCCCTGGTCTGATGGCTGTGGCTGCTGCTGCTGCTGCTGCTG 183  
Qy 423 AGCTGGGGAGTCTCGACACCTGCTGATGCTGTGCGACCCCGACCCAGCGGCTGTTGT 482  
Db 184 CGGTGGGGAGCCCTCGACCACTCCACGCTCTGCGACGCCAGCCAGGGCTGGTCTGC 243  
Qy 483 CAGCTGGGGAGTCTCGACACCTGCTGATGCTGTGCGACCCCGACCCAGCGGCTGTTGT 542  
Db 244 CAGCCGGGGAGCCCTCGACCACTCCACGCTCTGCGACGCCAGCCAGGGCTGGTCTGC 303  
Qy 543 AGCTGTGAGTGAATGCGCCAGCTGCTGATGCTGTGCGACCCCGACCCAGCGGCTGTTGT 602  
Db 304 AGCTGTGAGTGAATGCGCCAGCTGCTGATGCTGTGCGACCCCGACCCAGCGGCTGTTGT 363  
Qy 603 GTCTGTGCGCTGTGATGACGGTGGCTTACC 635  
Db 364 ATCCGCTGCCCTCGAGGAGCGGCTTACC 396

## RESULT 9

US-10-146-726-319  
; Sequence 319, Application US/10146726  
; Publication No. US20030129690A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Tamas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3330R1C308  
; CURRENT FILING DATE: 2002-05-15  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 319  
; LENGTH: 1266  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-146-726-319

Query Match 43.0%; Score 273; DB 14; Length 1266;  
Best Local Similarity 80.9%; Pred. No. 6.8e-69;  
Matches 318; Conservative 0; Mismatches 75; Indels 0; Gaps 0;  
Qy 243 GGTGACATGAGGGGAGCCCACTGATCCATCTTCTGGCCACTTCTTCTCTGCTCTC 302  
Db 4 GGGGACATGAGAGGACACCGAAGACCCACTCTGGCCCTTCTTCTCTGCTCTC 63  
Qy 303 TCAATGTTGTGTCGCCAGCTGTGCGGACACCCCTGTACCTGTCTTGGACACCACTCC 362  
Db 64 TCAAGTGGCTACCCAGCTGTGCGGACACCATGTACCTGCGCCCTCTGCTCTC 123  
Qy 363 TGCCACAGGGGTACCCCTGGTCTGATGGCTGTGGCTGCTGCTGCTGCTGCTGCTGCT 422

Db 124 TGCCCGCTGGAGTACCCCTGGTCTGATGCTGGCTGTGCTGCTGCTGCTGCTGCTGCTG 183  
Qy 423 AGCTGGGGAGTCTCGACACCTGCTGATGCTGTGCGACCCCGACCCAGCGGCTGTTGT 482  
Db 184 CGGTGGGGAGCCCTCGACCACTCCACGCTCTGCGACGCCAGCCAGGGCTGGTCTGC 243  
Qy 483 CAGCTGGGGAGTCTCGACACCTGCGCCAGTGGCGCCATGCGGCTGTGCTGCTTGGATGAGGATGACGGT 542  
Db 244 CAGCCGGGGAGCCCTCGACCACTCCACGCTCTGCGACGCCAGCCAGGGCTGGTCTGC 303  
Qy 543 AGCTGTGAGTGAATGCGCCAGCTGCTGATGCTGTGCGACCCCGACCCAGCGGCTGTTGT 602  
Db 304 AGCTGTGAGTGAATGCGCCAGCTGCTGATGCTGTGCGACCCCGACCCAGCGGCTGTTGT 363  
Qy 603 GTCTGTGCGCTGTGATGACGGTGGCTTACC 635  
Db 364 ATCCGCTGCCCTCGAGGAGCGGCTTACC 396

## RESULT 10

US-10-146-727-319  
; Sequence 319, Application US/10146727  
; Publication No. US20030129691A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tamas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3330R1C312  
; CURRENT FILING DATE: 2002-05-15  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 319  
; LENGTH: 1266  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-146-727-319

Query Match 43.0%; Score 273; DB 14; Length 1266;  
Best Local Similarity 80.9%; Pred. No. 6.8e-69;  
Matches 318; Conservative 0; Mismatches 75; Indels 0; Gaps 0;  
Qy 243 GGTGACATGAGGGGAGCCCACTGATCCATCTTCTGGCCACTTCTTCTCTGCTCTC 302  
Db 4 GGGGACATGAGAGGACACCGAAGACCCACTCTGGCCCTTCTTCTCTGCTCTC 63  
Qy 303 TCAATGTTGTGTCGCCAGCTGTGCGGACACCCCTGTACCTGTCTTGGACACCACTCC 362  
Db 64 TCAAGTGGCTACCCAGCTGTGCGGACACCATGTACCTGCGCCCTCTGCTCTC 123  
Qy 363 TGCCACAGGGGTACCCCTGGTCTGATGGCTGTGGCTGCTGCTGCTGCTGCTGCTGCT 422  
Db 124 TGCCCGCTGGAGTACCCCTGGTCTGATGGCTGTGGCTGCTGCTGCTGCTGCTGCTG 183  
Qy 423 AGCTGGGGAGTCTCGACACCTGCTGATGCTGTGCGACCCCGACCCAGCGGCTGTTGT 482  
Db 184 CGGTGGGGAGCCCTCGACCACTCCAGCTCTGCGACCGCCAGCGGCTGGTCTGCTGCT 243

QY 483 CAGCTGGGGCAGGCCCTGGCGGCATGGGCTGTGTCTCTTGGATGAGGATGACGGT 542  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 244 CAGCCGGGGCAGGACCGGTGGCGGGCCCTGTGCTCTTGGCAGAGCAGCAGC 303  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
QY 543 ACCTGTAGGTGAATGGCGGCAGGTACCTGGATGAGAGACCTTTAAACCCAAATTCGAGG 602  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 304 AGCTGTAGGTGAACGGCGCCCTGTATCGGGAAGGGAGACCTTCCAGCCCACTGCAGC 363  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
QY 603 GTCCTGTGCGCTGTGATGACGGTGGCTTCACC 635  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 364 ATCCGCTGCCGTGCGAGGACGGCGGCTTCACC 396  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

## RESULT 11

US-10-146-788-319  
; Sequence 319, Application US/10146788  
; Publication No. US20030129693A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P33301R1C322  
; CURRENT APPLICATION NUMBER: US/10/146,788  
; CURRENT FILING DATE: 2002-05-15  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 319  
; LENGTH: 1266  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-146-788-319

Query Match 43.0%; Score 273; DB 14; Length 1266;  
Best Local Similarity 80.9%; Pred. No. 6.8e-69;  
Matches 318; Conservative 0; Mismatches 75; Indels 0; Gaps 0;  
QY 243 GGTGACATGAGGGCAGCCACTGATCCATCTCTGGCCACTTCCTCTGCTGCTTCCTC 302  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 4 GGGACATGAGAGGCAGACACCCACCTCTCTGGCCCTCTCTCTCTCTCTCTCTCTCTC 63  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
QY 303 TCAATGGTGTGTCGCCAGCTGTGCGGACACCTCTACCTGTGCTGTGGACACACCCACG 362  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 64 TCAAGGTGCGTACCCAGCTGTGCGGACACCATGTACCTGCCCTTGGCCACTCCCGA 123  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
QY 363 TGCCACAGGGGGTACCCCTGTGCTGGATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 422  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 124 TGCCCGCTGGAGTACCCCTGTGCTGGATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 183  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
QY 423 AGGCTGGGGAGTCTCTCGACACCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 482  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 184 CGGCTGGGGAGTCTCTCGACACCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 243  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
QY 483 CAGCTGGGGAGGCGCCCTGGGCGCATGGGCTGTGTCTCTTGGATGAGGATGACGGT 542  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 244 CAGCCGGGGAGGCGCCCTGGGCGCATGGGCTGTGTCTCTTGGATGAGGATGACGGT 542  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
QY 543 CAGCTGGGGAGGCGCCCTGGGCGCATGGGCTGTGTCTCTTGGATGAGGATGACGGT 602  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 244 CAGCCGGGGAGGCGCCCTGGGCGCATGGGCTGTGTCTCTTGGATGAGGATGACGGT 602  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 543 AGCTGTGAGTGAATGGCGGCAGGTACCTGGATGAGAGACCTTTAAACCCAAATTCGAGG 602  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 304 ACCTGTAGGTGAACGGCGCCCTGTATCGGGAAGGGAGACCTTCCAGCCCACTGCAGC 363  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
QY 603 GTCCTGTGCGCTGTGATGACGGTGGCTTCACC 635  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 364 ATCCGCTGCCGTGCGAGGACGGCGGCTTCACC 396  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

## RESULT 12

US-10-152-380-319  
; Sequence 319, Application US/10152380  
; Publication No. US20030129694A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P33301R1C397  
; CURRENT APPLICATION NUMBER: US/10/152,380  
; CURRENT FILING DATE: 2002-05-21  
; Prior Application removed - See Palm or File Wrapper  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 319  
; LENGTH: 1266  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-152-380-319

Query Match 43.0%; Score 273; DB 14; Length 1266;  
Best Local Similarity 80.9%; Pred. No. 6.8e-69;  
Matches 318; Conservative 0; Mismatches 75; Indels 0; Gaps 0;  
QY 243 GGTGACATGAGGGCAGCCACTGATCCATCTCTGGCCACTTCCTCTGCTGCTTCCTC 302  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 4 GGGACATGAGAGGCAGACACCCACCTCTCTGGCCCTCTCTCTCTCTCTCTCTCTCTC 63  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
QY 303 TCAATGGTGTGTCGCCAGCTGTGCGGACACCTCTACCTGTGCTGTGGACACACCCACG 362  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 64 TCAAGGTGCGTACCCAGCTGTGCGGACACCATGTACCTGCCCTTGGCCACTCCCGA 123  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
QY 363 TGCCACAGGGGGTACCCCTGTGCTGGATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 422  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 124 TGCCCGCTGGAGTACCCCTGTGCTGGATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 183  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
QY 423 AGGCTGGGGAGTCTCTCGACACCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 482  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 184 CGGCTGGGGAGTCTCTCGACACCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 243  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
QY 483 CAGCTGGGGAGGCGCCCTGGGCGCATGGGCTGTGTCTCTTGGATGAGGATGACGGT 542  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 244 CAGCCGGGGAGGCGCCCTGGGCGCATGGGCTGTGTCTCTTGGATGAGGATGACGGT 542  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
QY 543 AGCTGTGAGTGAATGGCGGCAGGTACCTGGATGAGAGACCTTTAAACCCAAATTCGAGG 602  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 304 AGCTGTGAGTGAACGGCGCCCTGTATCGGGAAGGGAGACCTTCCAGCCCACTGCAGC 363  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
QY 603 GTCCTGTGCGCTGTGATGACGGTGGCTTCACC 635  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```
Db 364 ATCCGCTGCCGCTCGGAGGACGGCGCTTACC 396
||||| ||||| || ||||| ||||| |||||
RESULT 13
US-10-153-934-319
; Sequence 319, Application US/10153934
; Publication No. US20030129695A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang
; TITLE OF INVENTION:
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/028,072
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059836
; PRIOR FILING DATE: 1997-09-24
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062285
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062814
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/062816
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063045
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063082
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/063127
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063327
; PRIOR FILING DATE: 1997-10-27
; PRIOR APPLICATION NUMBER: 60/063329
; PRIOR FILING DATE: 1997-10-27
; PRIOR APPLICATION NUMBER: 60/063550
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063561
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063704
; PRIOR FILING DATE: 1997-10-29
;
Query Match 43.0%; Score 273; DB 14; Length 1266;
Best Local Similarity 80.9%; Pred. No. 6.8e-69;
Matches 318; Conservative 0; Mismatches 75; Indels 0; Gaps 0;
Qy 243 GGTGACATGAGGGGACGCCACTGATCCATCTTCTGGCCACTTCTTCCTCTGCTGCTCTC 302
|| ||||| ||||| || || ||||| || || ||||| || || ||||| || ||
Db 4 GGGACATGAGAGGACACACCGACACCCACTCTCTGCGCTTCTCCCTCTCTGCTCTCTC 63
Qy 303 TCAATGTGTGTGCGGAGCTGTGCGGACACCCCTGTACCTGTCTTGGACACCAACCACG 362
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 64 TCAAAGTGTGCTACCCAGCTGTGCGCGACACACCATGTACCTGCCCCCTGGCCACCTCCCCGA 123
Qy 363 TGCCACAGGGGGTACCCCTGGTGTGCTGGATGGCTGTGCTGCTAAAGTGTGTGACAGG 422
||||| || || ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 124 TGCCCGCTGGAGTACCCCTGGTGTGCTGGATGGCTGTGCTGCTGCGCGGTATGTGCACGG 183
Qy 423 AGGCTGGGGAGTCTCGGACCACTGTCATGTCTGCGACCCCGACCGGCGCTGGTTGT 482
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 184 CGGCTGGGGAGCCCTCGGACCACTCCAGCTCTGCGACGCGACGCGGCGCTGGTCTGC 243
Qy 483 CAGCTGGGCGAGCCCTGGCGGCATGGGCTGTGTCTCTTGGATGAGGATGACGGT 542
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 244 CAGCCCGGGGAGGACCCGGTGGCGGGGGGCCCTGTGCTCTTGGCAGAGGACGACAGC 303
Qy 543 AGCTGTGAGGTGAATGGCCGAGGTACCTGATGGAGAGACCTTTAAACCAATTCGAGG 602
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 304 AGCTGTGAGTGAACGCCCGCTGTATCGGAGGGGAGACCTTCCAGCCCACTGCAGC 363
Qy 603 GTCTGTGCCGCTGTGATGACGGTGGCTTACC 635
||| ||||| || ||||| || ||||| || ||||| || ||||| || |||||
Db 364 ATCCGCTGCCGCTCGGAGGACGGCGCTTACC 396
||||| ||||| || ||||| ||||| |||||
RESULT 14
```





```
Db 184 CGGCTGGGGAGCCCTCGGACCACTCCAGCTGCGACGCCAGCCAGGCGCTGGTCTGC 243
Qy 483 CAGCCTGGGGCAGGCCCTGGCGCCATGGGCTGTGTCTCTTGGATGAGGATGACGGT 542
Db 244 CAGCCCGGGCAGGACCCGGTGGCGGGGGCCCTGTGCTCTTGGCAGAGGACGACG 303
Qy 543 AGCTGTGAGTGAATGGCCGACGTACTCTGATGGAGAGACCTTTAAACCCCAATTGCAGG 602
Db 304 AGCTGTGAGTGAACGGCCGCTGTATCGGAAGGGGAGACCTTCCAGCCCCACTGCAGC 363
Qy 603 GTCTGTGCGCGCTGTGATGACGGTGGCTTACC 635
Db 364 ATCCGCTGCCGCTGCGAGGACGGCGGCTTACC 396
```

## RESULT 15

```
US-10-121-049-319
; Sequence 319, Application US/10121049
; Publication No. US20030022239A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C17
; CURRENT APPLICATION NUMBER: US/10/121,049
; CURRENT FILING DATE: 2002-04-12
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 319
; LENGTH: 1266
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-121-049-319
```

```
Query Match 43.0%; Score 273; DB 15; Length 1266;
Best Local Similarity 80.9%; Pred. No. 6.8e-69;
Matches 318; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

Qy 243 GGTGACATGAGGGGACGCCACCTGATCCATCTTTGGCCACTTCTCTTCCTTCGCTTCTC 302
Db 4 GGGGACATGAGGACACCCAGACCCACCTCTCGSCCTTCTCCCTCTCTCTGCTCTCTC 63

Qy 303 TCAATGTGTGTGCCAGCTGTGCCGGACACCCCTGTACCTGTCTTGGACACCCACCCAG 362
Db 64 TCAAAGGTGGGTACCCAGCTGTGCCCGACACCACTGTACCTGTGCCCTGGCCACCTTCCCGCA 123

Qy 363 TGCCACAGGGGTACCCCTGGTGTCTGGATGGCTGTGGCTGTCTAAAGTGTGTGCACGG 422
Db 124 TGCCCGTGGAGTACCCCTGGTGTCTGGATGGCTGTGGCTGTCTGGCTGTGTGCACGG 183

Qy 423 AGGCTGGGGAGTCTCGGACCACTGCTGTCTGCGACCCCGCCAGCCGCGCTGGTTGT 482
Db 184 CGGCTGGGGAGCCCTCGGACCACTCCAGCTGTGCGACCGCAGCCAGGCGCTGGTCTGC 243

Qy 483 CAGCCTGGGGCAGGCCCTGGCGCCATGGGCGCTGTGTCTCTTGGATGAGGATGACGGT 542
Db 184 CAGCCTGGGGCAGGCCCTGGCGCCATGGGCGCTGTGTCTCTTGGATGAGGATGACGGT 542
```

```
Db 244 CAGCCCGGGGCGAGGACCCCGTGGCCGGGGGCGCTGTGCTCTTGGCAGAGGACGACG 303
Qy 543 AGCTGTGAGTGAATGGCCCGCAGGTACTCTGATGGAGAGACCTTTAAACCCCAATTGCAGG 602
Db 304 AGCTGTGAGTGAACGGCCGCTGTATCGGAAGGGGAGACCTTCCAGCCCCACTGCAGC 363
Qy 603 GTCTGTGCGCGCTGTGATGACGGTGGCTTACC 635
Db 364 ATCCGCTGCCGCTGCGAGGACGGCGGCTTACC 396
```

Search completed: July 28, 2003, 21:31:54  
Job time : 157.533 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 18:02:17 ; Search time 1126 Seconds  
(without alignments)  
9133.322 Million cell updates/sec

Title: US-10-010-408-l\_COPY\_1\_635

Perfect score: 635

Sequence: 1 GAGGCTTCGATCTCCAGAG.....GTGATGACGGTGGCTTCACC 635

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

- 1: em\_estba:\*
- 2: em\_esthum:\*
- 3: em\_estin:\*
- 4: em\_estmu:\*
- 5: em\_estov:\*
- 6: em\_estpl:\*
- 7: em\_estro:\*
- 8: em\_hic:\*
- 9: gb\_estl:\*
- 10: gb\_est2:\*
- 11: gb\_hic:\*
- 12: gb\_est3:\*
- 13: gb\_est4:\*
- 14: gb\_est5:\*
- 15: em\_estfun:\*
- 16: em\_estom:\*
- 17: gb\_gss:\*
- 18: em\_gss\_hum:\*
- 19: em\_gss\_inv:\*
- 20: em\_gss\_pln:\*
- 21: em\_gss\_vrt:\*
- 22: em\_gss\_fun:\*
- 23: em\_gss\_mam:\*
- 24: em\_gss\_mus:\*
- 25: em\_gss\_other:\*
- 26: em\_gss\_pro:\*
- 27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|-------|---------------|--------|-------|-------------|
| 1          | 425.8 | 67.1          | 940    | 14    | BQ937887    |
| 2          | 385.2 | 60.7          | 537    | 14    | BQ560868    |
| 3          | 335.8 | 52.9          | 424    | 10    | BB849097    |
| 4          | 299   | 47.1          | 749    | 9     | AL555144    |
| 5          | 297.4 | 46.8          | 1166   | 13    | BM543799    |
| 6          | 271.4 | 42.7          | 1251   | 14    | BQ961357    |

|    |       |      |      |    |          |                    |
|----|-------|------|------|----|----------|--------------------|
| 7  | 269.4 | 42.4 | 452  | 10 | BE481184 | BE481184 166529 BA |
| 8  | 267.6 | 42.1 | 1058 | 14 | BM805088 | BM805088 AGENCOURT |
| 9  | 267.2 | 42.1 | 620  | 13 | BG928868 | BG928868 HNC57-1-D |
| 10 | 262   | 41.3 | 750  | 13 | BM043988 | BM043988 603620978 |
| 11 | 261.6 | 41.2 | 1006 | 14 | BM921531 | BM921531 AGENCOURT |
| 12 | 258.2 | 40.7 | 928  | 13 | BI161474 | BI161474 602864871 |
| 13 | 258.2 | 40.7 | 979  | 14 | BQ279131 | BQ279131 AGENCOURT |
| 14 | 256.8 | 40.4 | 933  | 14 | BQ278961 | BQ278961 AGENCOURT |
| 15 | 256   | 40.3 | 1073 | 14 | BQ073722 | BQ073722 AGENCOURT |
| 16 | 247.8 | 39.0 | 800  | 13 | BI826781 | BI826781 603077268 |
| 17 | 246.2 | 38.8 | 888  | 13 | BI825652 | BI825652 603072631 |
| 18 | 245   | 38.6 | 886  | 13 | BI822142 | BI822142 603039845 |
| 19 | 238.2 | 37.5 | 651  | 13 | BI457141 | BI457141 603185392 |
| 20 | 229   | 36.1 | 916  | 12 | BI457367 | BI457367 603185689 |
| 21 | 223.2 | 35.1 | 405  | 12 | BG900069 | BG900069 HOA51-1-A |
| 22 | 209.6 | 33.0 | 380  | 12 | BG900020 | BG900020 HOA48-1-C |
| 23 | 205.8 | 32.4 | 618  | 12 | BG538695 | BG538695 602566932 |
| 24 | 205.2 | 32.3 | 792  | 13 | BI823598 | BI823598 603040962 |
| 25 | 161.2 | 25.4 | 657  | 13 | BM488499 | BM488499 p9m2b.pk0 |
| 26 | 145.4 | 22.9 | 790  | 13 | BM046275 | BM046275 603626068 |
| 27 | 144.4 | 22.7 | 1022 | 14 | BQ952960 | BQ952960 AGENCOURT |
| 28 | 117.6 | 18.5 | 742  | 13 | BI758148 | BI758148 603023866 |
| 29 | 110.2 | 17.4 | 190  | 9  | AA647775 | AA647775 vp04c09.r |
| 30 | 92.8  | 14.6 | 502  | 10 | AV670831 | AV670831 AV670831  |
| 31 | 87.2  | 13.7 | 828  | 9  | AL543019 | AL543019 AL543019  |
| 32 | 85.6  | 13.5 | 505  | 12 | BF037009 | BF037009 601456766 |
| 33 | 85.6  | 13.5 | 622  | 14 | BM772255 | BM772255 K-EST0056 |
| 34 | 85.6  | 13.5 | 756  | 9  | AU125723 | AU125723 AU125723  |
| 35 | 85.6  | 13.5 | 826  | 12 | BG116760 | BG116760 602318826 |
| 36 | 85.6  | 13.5 | 827  | 9  | AU131392 | AU131392 AU131392  |
| 37 | 85.6  | 13.5 | 851  | 13 | BI828218 | BI828218 603073692 |
| 38 | 85.6  | 13.5 | 874  | 14 | BQ720581 | BQ720581 AGENCOURT |
| 39 | 85.6  | 13.5 | 877  | 9  | AL546864 | AL546864 AL546864  |
| 40 | 85.6  | 13.5 | 878  | 14 | BQ719477 | BQ719477 AGENCOURT |
| 41 | 85.6  | 13.5 | 887  | 14 | BQ230898 | BQ230898 AGENCOURT |
| 42 | 85.6  | 13.5 | 888  | 9  | AL544380 | AL544380 AL544380  |
| 43 | 85.6  | 13.5 | 896  | 14 | BQ719732 | BQ719732 AGENCOURT |
| 44 | 85.6  | 13.5 | 931  | 14 | BQ880701 | BQ880701 AGENCOURT |
| 45 | 85.6  | 13.5 | 942  | 9  | AL544453 | AL544453 AL544453  |

# ALIGNMENTS

RESULT 1  
BQ937887

LOCUS

DEFINITION AGENCOURT\_8951807 NCI\_CGAP\_Co24 Mus musculus cdna clone  
IMAGE:6476852 5', mRNA sequence.

ACCESSION BQ937887

VERSION BQ937887.1

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 940)

AUTHORS

TITLE

JOURNAL

COMMENT

NIH-MGC http://mgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: The Cepko Laboratory

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM14017 row: n column: 21

High quality sequence stop: 543.

Location/Qualifiers

1. 940

```
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:6476852"
/clone_lib="NCI CGAP Co24"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.6 kb. Constructed by Life
Technologies. Note: this is a NCI CGAP Library."

BASE COUNT      169 a  277 c  288 g  200 t      6 Others
ORIGIN

Query Match      67.1%; Score 425.8; DB 14; Length 940;
Best Local Similarity 90.9%; Pred. No. 1.5e-99;
Matches 488; Conservative 0; Mismatches 39; Indels 10; Gaps 3;

QY 1 GACGGTTCGTATCCAGAGACCCCTGGGGTGGGACAGGGGCGCTTGGCAAGCGTCGACCC 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 37 GACGCTCCTGTATCTCCAGAGACCCCGGGCTGGGACAGGGGCGCTTGGGAGCGTCGACGT 96
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 61 GCTG-GGCAGTGGCTTGAATGAGGCTCTTTATTACTGGGAACCTGAGGAGCTAAGAGCT 119
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 97 GCTGTGGCAGTAGCTTGGGATGAGGCTCTTTCTGCTGGGAACCTGAGGAGCTGAGAGCT 156
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 120 CTTGTCAG---CTTGCTCTAAAGTCTTAGCACTTGTGTGCTTGGGCTTCACACACTGT 176
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 157 CTTGTCAGGCTCTGTCTTAAACTCTTGGCACTTGGGCTTGGGCTTCACACACTGT 216
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 177 CAGACACTTGTGGTGGCTCCAGGCTCACCTTCACTTCACTTGAAGTGTGCTCCACAAG 236
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 217 CAGACACTTCTTGGTGCCCTCTCGGCC-----TCAGGTTTGAAGTGTGCTCCACAAG 270
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 237 GGACACGCTGACATGAGGGGAGCCCACTGATCCATCTTCTGGCCACTTCTTCTCTGCG 296
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 271 GGACACGCTGACATGAGGGGAGCCCACTGATCCATCTTCTGGCCACTTCTTCTCTGCG 330
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 297 CTTCTCTCAATGGTGTGCCAGCTGTGCCGACACCCCTGTACTGTCTTGGACACCA 356
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 331 ATTCTCTCAATGGTGTATGCCAGCTGTGCCAGACACCCCTGTGCTCTTGGACACCA 390
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 357 CCCAGTCCACAGGGGATACCCTGGTGTGGATGCTGCTGCTGTGTAAGTGTGT 416
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 391 CCCAGTCCACAGGGGATACCCTGGTGTGGATGCTGCTGCTGTGTAAGTGTGT 450
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 417 GCACGAGGCTGGGGAGTCTTGGACACCTGTGCTGCGACCCCGACGAGGCGCTG 476
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 451 GCACGAGGCTGGGGAGTCTTGGACACCTGTGCTGCGACCCCGACGAGGCGCTG 510
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 477 GTTGTGACCTGGGGAGGCGCTTGGGCGCATGGGCTGTGTCTTGGATGAG 533
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 511 GTTGTGACCTGGGGAGGCGCTTGGGCGCATGGGCTGTGTCTTGGATGAG 567
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 2
BQ560868
LOCUS
DEFINITION
H4067A01-5 NIA Mouse 7.4K cDNA Clone Set Mus musculus cDNA clone
H4067A01 5', mRNA sequence.
ACCESSION
BQ560868
VERSION
BQ560868.1 GI:214161753
KEYWORDS
EST.
SOURCE
house mouse.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 537)
AUTHORS
VanBuren,V., Piao,Y., Dudekula,D.B., Qian,Y., Carter,M.G., Martin
P.R., Stagg,C.A., Bassey,U., Aliba,K., Hamatani,T., Kargul,G.J.,
Luo,A.G. and Ko,M.S.H.
TITLE
Assembly, verification, and initial annotation of NIA 7.4K mouse
cDNA clone set
JOURNAL
Unpublished (2002)

COMMENT
Contact: Yong Qian
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 3000, Baltimore, MD 21224-6820, USA
Email: cdna@igsun.grc.nia.nih.gov
This clone set has been freely distributed to the community. Please
visit http://igsun.grc.nia.nih.gov/cDNA/NIA_7.4k.html for details.
Plate: H4067 Row: A Column: 01
Seq primer: -20M13 Reverse
High quality sequence stop: 537
POLYA-No.

FEATURES
Location/Qualifiers
1..537
/organism="Mus musculus"
/strain="C57BL/6"
/db_xref="niaEST:H4067A01-5"
/db_xref="taxon:10090"
/clone="H4067A01"
/clone_lib="NIA Mouse 7.4K cDNA Clone Set"
/sex="mixed"
/dev_stage="mixed"
/lab_host="DH10B"
/note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; This
clone is among a rearranged set of 7,407 clones from more
than 20 cDNA libraries."
BASE COUNT      87 a  162 c  166 g  122 t

Query Match      60.7%; Score 385.2; DB 14; Length 537;
Best Local Similarity 92.1%; Pred. No. 3.7e-89;
Matches 442; Conservative 0; Mismatches 28; Indels 10; Gaps 3;

QY 1 GACGCTTCGTATCTCCAGAGACCCCTGGGGTGGGACAGGGGCGCTTGGCAAGCGTCGACCC 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 64 GACGCTTCGTATCTCCAGAGACCCCGGGCTGGGACAGGGGCGCTTGGGAGCGTCGACCT 123
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 61 GCTG-GGCAGTGGCTTGAATGAGGCTCTTTATTACTGGGAACCTGAGGAGCTAAGAGCT 119
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 124 GCTGTGGCAGTAGCTTGGGATGAGGCTCTTTCTGCTGGGAACCTGAGGAGCTGAGAGCT 183
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 120 CTTGTCAG---CTTGCTCTAAAGTCTTAGCACTTGTGTGCTTGGGCTTCACACACTGT 176
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 184 CTTGTCAGGCTGCTGTCTTAACTTGTGGCACTTGGGCTGCTTGGGCTTCACACACTGT 243
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 177 CAGACACTTCTGCTGGGCTCCAGGCTCACCTTCACTTCAAGTTTGAAGTGTGCTCCACAAG 236
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 244 CAGACACTTCTGCTGGGCTCCCTCGGCC-----TCAGGTTTGAAGTGTGCTCCACAAG 297
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 237 GGACACGCTGACATGAGGGGAGCCCACTGATCCATCTTCTGGCCACTTCTTCTCTGCG 296
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 298 GGACACGCTGACATGAGGGGAGCCCACTGATCCATCTTCTGGCCACTTCTTCTCTGCG 357
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 297 CTTCTCTCAATGGTGTGTGCCAGCTGTGCCGACACCCCTGTACTGTCTTGGACACCA 356
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 358 ATTCTCTCAATGGTGTATGCCAGCTGTGCCAGACCCCTGTGCTGCTTGGACACCA 417
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 357 CCCAGTCCACAGGGGATACCCTGGTGTGGATGCTGCTGCTGTGTAAGTGTGT 416
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 418 CCCAGTCCACAGGGGATACCCTGGTGTGGATGCTGCTGCTGTGTAAGTGTGT 477
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 417 GCACGAGGCTGGGGAGTCTTGGACACCTGTGCTGCGACCCCGACGAGGCGCTG 476
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 478 GCACGAGGCTGGGGAGTCTTGGACACCTGTGCTGCGACCCCGACGAGGCGCTG 537
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 3
BB849097
LOCUS
DEFINITION
BB849097 RIKEN full-length enriched, adult inner ear Mus musculus
cDNA clone F93006G02 5', mRNA sequence.
ACCESSION
BB849097
VERSION
BB849097.1 GI:17090551
KEYWORDS
EST.

BB849097 424 bp mRNA linear EST 26-NOV-2001
BB849097 RIKEN full-length enriched, adult inner ear Mus musculus
cDNA clone F93006G02 5', mRNA sequence.
```

SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
house mouse.  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 424)  
Akimura,T., Arawaka,T., Carninci,P., Furuno,M., Hanagaki,T.,  
Hayatsu,N., Hiramoto,K., Hirooka,T., Hirozane,T., Imotani,K., Ishii  
Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Matsuyama,T.,  
Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T.,  
Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K.,  
Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa  
A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T.,  
Watanishi,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.  
RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.  
2001)  
Unpublished (2001)  
Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@gsc.riken.go.jp,  
URL: http://genome.gsc.riken.go.jp/  
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh  
M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new  
genes. Genome Res. 10 (10), 1617-1630 (2000)  
wagi,K., Fujiwaki,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,  
Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura  
S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and  
Hayashizaki,Y.  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
10 (11), 1757-1771 (2000)  
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara  
Y. and Hayashizaki,Y.  
Computer-based methods for the mouse full-length cDNA  
encyclopedia: real-time sequence clustering for construction of a  
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
Please visit our web site (http://genome.gsc.riken.go.jp) for  
further details.  
e mouse tissues.  
FEATURES  
source  
1. 424  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="F030006G02"  
/clone\_lib="RIKEN full-length enriched, adult inner ear"  
/tissue\_type="inner ear"  
/dev\_stage="adult"  
/note="pooled tissues ; (tissue\_type=cerebellum,  
dev\_stage=16 days neonate, sex=mixed),  
(tissue\_type=cerebellum, dev\_stage=0 day neonate,  
sex=mixed), (tissue\_type=hippocampus, dev\_stage=adult,  
sex=adult), (tissue\_type=whole body, dev\_stage=9 days  
embryo, sex=mixed), (tissue\_type=lung, dev\_stage=13 days  
embryo, sex=mixed)"  
BASE COUNT 65 a 124 c 132 g 103 t  
ORIGIN  
Query Match 52.9%; Score 335.8; DB 10; Length 424;  
Best Local Similarity 91.4%; Pred. No. 2e-76;  
Matches 392; Conservative 0; Mismatches 27; Indels 10; Gaps 3;  
QY 37 AGGGCCCTTGCAAGGCTGCAGCGCTG-GGCAGTGGCTTGAATGGAGGCTTTATTAC 95  
|||||  
DB 2 AGGGCCCTTGCGAGGCTGCAGCTGGTGGCAGTAGCTGGGATGGAGGCTCTTCTGC 61  
|||||  
QY 96 TGGGAACCTGAGGAGCTAAGAGGCTCCTGTCAG-CTTGTCCTAAAGTCTTAGCACTTGT 152  
|||||

Db 62 TGGGAACCTGAGGAGCTGAGAAGCTCTCTGTGAGGCTCTCTAAACTCTTGGCAGCTGC 121  
|||||  
QY 153 GGTGGCTTGGCTTCACACACTGTTCAGACACCTCTCGTGGTGGCTCCACGCGCTCACCTT 212  
|||||  
Db 122 GGTGGCTTGGCTTCACACACTGTTCAGACACCTCTCTGTGGTGGCTCTCGGCC- -T 175  
|||||  
QY 213 CAGGTTTGAAGCTGGCTCCACAAGGACACGGTGACATGAGGGGACGCCCACTGATCCAT 272  
|||||  
Db 176 CAGGTTTGAAGCTGGCTCCACAAGGACACGGTGACATGAGGGGACGCCCACTGATCCAT 235  
|||||  
QY 273 CTCTTGCCACTTCT 332  
|||||  
Db 236 CTCTTGCCACTTCT 295  
|||||  
QY 333 CCCTGTACCTGTCTCTTGACACACCCAGTGGCCACAGGGGTACCCCTGGTCTGCTGCAT 392  
|||||  
Db 296 CCCTGTGCTGTCTCTTGACACACCCAGTGGCCACAGGGGTACCCCTGGTCTGCTGCAT 355  
|||||  
QY 393 GGCTGTGGCTGCTGTAAAGTGTGTGACAGGAGCTGGGGAGTCTCTCGACCACTGCAT 452  
|||||  
Db 356 GGCTGTGGCTGCTGTGAGTGTGTGACAGGAGCTGGGGAGTCTCTCGACCACTGCAT 415  
|||||  
QY 453 GTCTGCGAC 461  
|||||  
Db 416 GTCTGCAAC 424  
|||||  
RESULT 4  
AL555144 749 bp mRNA linear EST 16-FEB-2001  
LOCUS LTI\_NFL006.PL2 Homo sapiens cDNA clone CS0DK007Y021 5  
DEFINITION prime, mRNA sequence.  
ACCESSION AL555144  
VERSION AL555144.1 GI:12896595  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 749).  
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.  
FEATURES  
source  
1. 749  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="CS0DK007Y021"  
/clone\_lib="LTI\_NFL006.PL2"  
/tissue\_type="placenta"  
/note="vector: pCMVSPORT 6; Site\_1: NotI; 1st strand cDNA  
was primed with a NotI-oligo(GT) primer. Five prime end  
enriched, double-stranded cDNA was digested with Not I and  
cloned into the Not I and Eco RV sites of the pCMVSPORT 6  
vector. Library was normalized. Library was constructed by  
Life Technologies. Contact : Feng Liang Life Technologies,  
a division of Invitrogen 9800 Medical Center Drive  
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371  
Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com"  
BASE COUNT 128 a 252 c 236 g 133 t  
ORIGIN  
Query Match 47.1%; Score 299; DB 9; Length 749;  
Best Local Similarity 79.1%; Pred. No. 8.7e-67;  
Matches 371; Conservative 0; Mismatches 90; Indels 8; Gaps 1;  
QY 167 CACACACTGTACAGACACCTTCGTGGTGGCTCCACGGCCTCACCTCAGGTTTGAAGCTG 226  
|||||



```

/clone="IMAGE:6484532"
/clone_lib="NIH_MGC_40"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
178 a 393 c 546 g 131 t 3 others
BASE COUNT

```

| Query Match           | 42.7% | Score  | 271.4   | DB         | 14 | Length | 1251 |
|-----------------------|-------|--|---------|------------|----|--------|------|
| Best Local Similarity | 80.7% | Pred. No.  | 1.5e-59 |            |    |        |      |
| Matches               | 317   | Conservative   | 0       | Mismatches | 76 | Indels | 0    |
| QY                    | 243   | GGTGACATGAGGGCAGCCACCTGATCCATCTCTGGCCACTTCCTTCCTCCCTGCCCTCTC | 302     |            |    |        |      |
| DB                    |       |  |         |            |    |        |      |
| DB                    | 8     | GGGACATGAGGGCACACCGAAGACCCACTCTGGCCCTTCCTCCCTTCCTGCCCTCTC    | 67      |            |    |        |      |
| QY                    | 303   | TCAATGTGTGTGCCAGCTGTGCCGACACCCCTGTACCTGTCTTGGACACACACCCACG   | 362     |            |    |        |      |
| DB                    |       |  |         |            |    |        |      |
| DB                    | 68    | TCAAAGTGGCTACCCAGCTGTCCGACACCATGTACCTGCCCTGGCCACCTCCCCGA     | 127     |            |    |        |      |
| QY                    | 363   | TGCCACACGGGGTACCCCTGTGTCTGGATGGCTGTGGCTGTGAAGTGTGTGCACGG     | 422     |            |    |        |      |
| DB                    |       |  |         |            |    |        |      |
| DB                    | 128   | TGCCCGCTGGGAGTACCCCTGGTGTGTGATGGCTGTGGCTGTGCCGGGTATGTGCACGG  | 187     |            |    |        |      |
| QY                    | 423   | AGGCTGGGGAGTCTTCGACACCACTGCATGTCTTGCCACCCAGCACAGGSCCTGGTTTGT | 482     |            |    |        |      |
| DB                    |       |  |         |            |    |        |      |
| DB                    | 188   | CGGCTGGGGAGCCCTCGACCAACTCCACGTCTGCGACCGACCGAGGCCCTGGCTGC     | 247     |            |    |        |      |
| QY                    | 483   | CAGCCTGGGCAGGCCCTTGGCGGCCATGCGGCTGTGTGTCTTGTGATGAGGATGACGGT  | 542     |            |    |        |      |
| DB                    |       |  |         |            |    |        |      |
| DB                    | 248   | CAGCCCGGGCAGGACCCCGTGGACGGGGGGCCCTGTGCCTCTTGGCAGAGACGACGC    | 307     |            |    |        |      |
| QY                    | 543   | AGCTGTGAGTGAATGGCCGACGTACTCGATGGAGAGACCTTTAAACCCCAATTGCAGG   | 602     |            |    |        |      |
| DB                    |       |  |         |            |    |        |      |
| DB                    | 308   | AGCTGTGAGTGAACGGCCGCTGTATCGGGAGGGGAGACCTTCCAGCCCCACTGCAGC    | 367     |            |    |        |      |
| QY                    | 603   | GTCCGTGGCCGTGTGATGACGGTGGCTTCACC                             | 635     |            |    |        |      |
| DB                    |       |  |         |            |    |        |      |
| DB                    | 368   | ATCCGCTGCCGCTGCGAGGACGGCGGCTTCACC                            | 400     |            |    |        |      |

[illegible]

REFERENCE  
AUTHORS Bovidae; Bovinae; Bos.  
1 (bases 1 to 452)  
Sonstegard, R.S., Capuco, A.V., Van Tassell, C.P., Ashwell, M.S. and  
Wells, K.D.  
TITLE Mapping of Expressed Sequence Tags from a normalized bovine mammary  
gland cDNA library  
JOURNAL Unpublished (2000)  
COMMENT Contact: Sonstegard TS  
USDA, ARS, Beltsville Agricultural Research Center  
Bldg. 200 Rm 2A, Beltsville, MD 20705, USA  
Tel: 301 504 8416  
Fax: 301 504 8414  
Email: tads@anri.barc.usda.gov  
Single pass sequencing. Bases called and alt trimmed with phred

```

v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.
PCR Primers
FORWARD: AGGAAACAGCTATGACCAT
BACKWARD: GTTTTCCGAGTCACGACG
Plate: 9 row: E column: 17
Seq primer: ATTTAGGTGACACTATAG.
Location/Qualifiers
1. .452
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_lib="BARC 5BOV"
/tissue_type="pooled"
/lab_host="DH10B"
/note="vector: pCMV SPORT6; Site.1: NotI; Site.2: SalI;
Library made from pooled mRNA isolated from mammary
tissues at eight physiological, developmental, and disease
states."
59 a 158 c 152 g 83 t
BASE COUNT

```

|                       |       |   |         |            |    |        |     |
|-----------------------|-------|---|---------|------------|----|--------|-----|
| Query Match           | 42.4% | Score   | 269.4   | DB         | 10 | Length | 452 |
| Best Local Similarity | 79.7% | Pred. No.   | 3.2e-59 |            |    |        |     |
| Matches               | 318   | Conservative  | 0       | Mismatches | 81 | Indels | 0   |
| Gaps                  | 0     |   |         |            |    |        |     |
| QY                    | 236   | GGGACAGGTGACATGAGGGGACGCCACATGATCCATCTTCGCGCACATTCCTTCCTCTG | 295     |            |    |        |     |
| Db                    | 5     | GGGACAGGTGACATGAGGGGACATTTGAGACCCACCTCCTCGCCTTCCTCCCTCTCTG  | 64      |            |    |        |     |
| QY                    | 296   | CTTCTCTCAATGGTGTGCCAGCTGTGCCGGACACCCGTGACCTGTCTTTGGACACC    | 355     |            |    |        |     |
| Db                    | 65    | TCCTCTCCAAAGTGTGTGCCAGCTGTGCCCCACACACCATGTGCCTGTCCCTGGCAC   | 124     |            |    |        |     |
| QY                    | 356   | ACCCAGTGCACACAGGGGTACCCCTGTGCTGGATGGCTGTGGCTGTCTGTAAGTGTG   | 415     |            |    |        |     |
| Db                    | 125   | ACCCGATGCCACAGGGGTGCCCCGTGGTGTGGACGGTGCACCTTCCTGCCGGGTATG   | 184     |            |    |        |     |
| QY                    | 416   | TGCACGAGGCTGGGGAGTCTTCGACACACCTTGCATGTCTGGACCCACGACGGGCCT   | 475     |            |    |        |     |
| Db                    | 185   | TGCGGCGGCTGGGGAGCCCTGCGACCATCTCCAGCTGTGTGACCCACGACGGCCT     | 244     |            |    |        |     |
| QY                    | 476   | GGTTTGTACGCTGGGGCAGGCCCTGGCGCCCATGGGCTGTGTCTTGTGATGAGGA     | 535     |            |    |        |     |
| Db                    | 245   | GGTCTGCCACCCGGGGGGGCCCTGGAGCCGGGGGGCGTGTGTCTCTGGGGAGAGA     | 304     |            |    |        |     |
| QY                    | 536   | TGACGGTGTGAGTGAATGGCCGACAGTACCTGGATGGAGAGACCTTTTAAACCCAA    | 595     |            |    |        |     |
| Db                    | 305   | TGACGCGAGCTGGAGGTGAACGCCCTCTACCGGGACGGGAGACCTTTTCAGCCTCA    | 364     |            |    |        |     |
| QY                    | 596   | TTGACAGGTCCTGTGGCGTGTGATGACGGTGGCTTAC                       | 634     |            |    |        |     |
| Db                    | 365   | CTGACAGCTCCGTCGCGTTCGAGACGACGGCGCTTAC                       | 403     |            |    |        |     |

|            |  |             |              |                          |  |  |  |  |  |
|------------|--|-------------|--------------|--------------------------|--|--|--|--|--|
| RESULT 8   | BM805088   | 1058 bp     | linear       | EST 05-MAR-2002          |  |  |  |  |  |
| LOCUS      | ACGNCQCOURT_6490429  | NIH_MGC_125 | Homo sapiens | cdna clone IMAGE:5725937 |  |  |  |  |  |
| DEFINITION | 5', mRNA sequence.   |             |              |                          |  |  |  |  |  |
| ACCESSION  | BM805088   |             |              |                          |  |  |  |  |  |
| VERSION    | BM805088.1   | GI:19121911 |              |                          |  |  |  |  |  |
| KEYWORDS   | EST.   |             |              |                          |  |  |  |  |  |
| SOURCE     | human.   |             |              |                          |  |  |  |  |  |
| ORGANISM   | Homo sapiens   |             |              |                          |  |  |  |  |  |
| REFERENCE  | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.                   |             |              |                          |  |  |  |  |  |
| AUTHORS    | NTH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .  |             |              |                          |  |  |  |  |  |
| TITLE      | National Institutes of Health, Mammalian Gene Collection (MGC)   |             |              |                          |  |  |  |  |  |
| JOURNAL    | Unpublished (1999)   |             |              |                          |  |  |  |  |  |
| COMMENT    | Contact: Robert Strausberg, ph.D.<br>Email: <a href="mailto:cgapbs-r@mail.nih.gov">cgapbs-r@mail.nih.gov</a><br>Tissue Procurement: Invitrogen |             |              |                          |  |  |  |  |  |

Kumar, S., Connor, J.R., Dodds, R.A., Halsey, W., Van Horn, M., Mao, J.,  
 Sathe, G., Mul, P., Agarwal, P., Badger, A.M., Lee, J.C., Gowen, M. and  
 Lark, M.W.  
 Identification and initial characterization of 5000 expressed  
 sequenced tags. (ESTs) each from adult human normal and  
 osteoarthritic cartilage cDNA libraries  
 Osteoarthr. Cartil. 9 (7), 641-653 (2001)  
 21482651  
 Contact: Sanjay Kumar  
 UW2109  
 GlaxoSmithKline  
 709 Swedeland Road, P.O. Box 1539, King of Prussia, PA 19406, USA  
 Tel: 610-270-7245  
 Fax: 610-270-5598  
 Email: sanjay\_kumar-1@gsk.com  
 Seq primer: T7  
 Location/Qualifiers  
 1. 620  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="HNC (Human Normal Cartilage)"  
 /tissue\_type="cartilage"  
 /lab\_host="E.coli DH10 B"  
 /note="Vector: pSPORT I; Site\_1: SalI; Site\_2: NotI;  
 Directional"

[illegible]

```

RESULT 10
BM043988
LOCUS
DEFINITION
750 bp linear mRNA
BM043988 Homo sapiens CDNA clone IMAGE:5446794 5',
mRNA sequence.
ACCESSION
BM043988
VERSION
BM043988.1 GI:16773255
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```



REFERENCE 1 (bases 1 to 750)  
NIH-MGC http://mgc.nci.nih.gov/  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
TITLE Unpublished (1999)  
JOURNAL  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabps-r@mail.nih.gov  
Tissue Procurement: DCTD/DTF  
cDNA Library Preparation: Ling Hong/Rubin Laboratory  
National Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
JOURNAL  
COMMENT Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: L1CMI929 row: k column: 19  
High quality sequence stop: 714.

FEATURES  
Source  
1..750  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5446794"  
/clone\_lib="NIH\_MGC\_40"  
/tissue\_type="carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: prostate; Vector: pOTB7; Site\_1: XhoI; Site\_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAGGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."  
Note: this is a NIH\_MGC Library."

BASE COUNT 112 a 267 c 246 g 125 t

Query Match 41.3%; Score 262; DB 13; Length 750;  
Best Local Similarity 80.7%; Pred. No. 3.3e-57;  
Matches 318; Conservative 0; Mismatches 75; Indels 1; Gaps 1;

Qy 243 GGTGACATGAGGGGAGCCCACTGATCATCTTCTGGCCACTTCTCTCCCTGCTGCTTC 302  
Db 10 GGTGACATGAGGGGAGCCCACTGATCATCTTCTGGCCACTTCTCTCCCTGCTGCTTC 69

Qy 303 TCATGTTGTGTCGCCAGCTGTGCGGACACCCCTGTACCTGTCTTGGACACACCCAG 362  
Db 70 TCAAAGTGTGTCGCCAGCTGTGCGGACACCCCTGTACCTGTCTTGGACACCCAG 129

Qy 363 TGCCACAGGGGTACCCCTGGTCTGATGGCTGTGCTGCTCTAAAGTGTGTGCACGG 422  
Db 130 TGCCCGCTGGGAGTACCCCTGGTGTGATGGCTGTGCTGCTGCTGCTGCTGCTGCTG 189

Qy 423 AGGCTGGGGAGTCTCTGCGACACCTGATGCTCTGCGACCCCA-GCCAGGGCCCTGGTTG 481  
Db 190 CGGCTGGGGAGCCCTGGGACCACTCCAGCTCTGCGACGCCATGCGAGGCCCTGGTGTG 249

Qy 482 TCAGCTGGGGAGCCCTGGCGCCATGGGGCTGTGTGCTCTTTGGATGAGGATGACGG 541  
Db 250 CCAGCCCGGGGAGCCCGGTGGACGGGGGGCCCTGTGCTCTTGGCAGAGGACGACAG 309

Qy 542 TAGCTGTGAGTGAATGCGCCAGCTACCTGGATGGAGAGACCTTTAAACCCATTGCG 601  
Db 310 CAGCTGTGAGTGAACCGCCCTGTATCGGGAGGGGAGACCTTCCAGCCCCACTGCG 369

Qy 602 GGTCCTGTGCGCTGTGATGACGGTGGCTTCACC 635  
Db 370 CATCCGCTGCGCTGCGGAGGAGCGGCGCTTCACC 403

RESULT 11  
BM921531 1006 bp mRNA linear EST 12-MAR-2002  
LOCUS  
DEFINITION AGENCOURT\_6708025 NIH\_MGC\_115 Homo sapiens cDNA clone IMAGE:5753009  
5', mRNA sequence.  
ACCESSION  
BM921531

BM921531.1 GI:19371910  
EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1006)  
NIH-MGC http://mgc.nci.nih.gov/  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
TITLE Unpublished (1999)  
JOURNAL  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabps-r@mail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: L1AMI2788 row: b column: 18  
High quality sequence stop: 593.

FEATURES  
Source  
1..1006  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5753009"  
/clone\_lib="NIH\_MGC\_115"  
/lab\_host="DH10B"  
/note="Organ: pooled brain, lung, testis; Vector: pCMV-SPORT6; Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH\_MGC Library."

BASE COUNT 148 a 368 c 317 g 169 t 4 others

ORIGIN

Query Match 41.2%; Score 261.6; DB 14; Length 1006;  
Best Local Similarity 80.5%; Pred. No. 4.7e-57;  
Matches 306; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

Qy 256 GCAGCCACATGATCATCTTCTGGCCACTTCTCTCCCTGCTGCTCTCAATGCTGTG 315  
Db 1 GCACACGAGAGACCCACTCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 60

Qy 316 CCAGCTGTGCGGACACCCCTGTACCTGTCTTGGACACACCCAGTGCACAGGGG 375  
Db 61 CCAGCTGTGCGGACACCATGTACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120

Qy 376 TACCCCTGGTGTGGATGGCTGTGGCTGTGTAAGTGTGTGCACGAGGCTGGGGAGT 435  
Db 121 TACCCCTGGTGTGGATGGCTGTGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180

Qy 436 CTTGCGACACCTGCTGCTGCGACCCAGCCAGGCGCTGTTGTTCAGCTGGGCGAG 495  
Db 181 CTTGCGACCACTCCAGCTGTGCGACCCAGCCAGGCGCTGCTGCTGCTGCTGCTGCTG 240

Qy 496 GCGCTGTGCGGCGCTGTGCTGTCTCTTGGATGAGGATGAGGTAGCTGTGAGTGA 555  
Db 241 GACCCGCTGGAGCGGGGCCCTGTGCTCTTGGCAGAGGACGACAGCAGTGTGAGTGA 300

Qy 556 ATGCCCGCAGGTACCTGGATGGAGAGACCTTTAAACCAATTGACGGGCTCCTGTGCCGCT 615  
Db 301 ACGCCCGCTGTATCGGGAAGGGAGACCTTCCAGCCCCACTGACGATCCGCTGCCGCT 360

Qy 616 GTGATGACGCTGGCTTCACC 635  
Db 361 GCGAGGACGCGGCTTCACC 380

|                           |   |   |     |
|---------------------------|---|---|-----|
| Qy                        | 569   | CCTCGATGGAGAGACACTTTAAACCATTTCGAGGGTCTCTGTGGCCTGTGATGACGGTGG      | 628 |
| Db                        | 382   | TCCGGAAGGGAGACCTTCAGCCCCACTGCAGCATCCGCTGCCGTGCGAGGACGCGG          | 441 |
| Qy                        | 629   | CTTCACC   | 635 |
| Db                        | 442   | CTTCACC   | 448 |
| RESULT 13                 |   |   |     |
| BQ279131                  |   |   |     |
| LOCUS                     | BQ279131  | 979 bp mRNA linear EST 07-MAY-2002                                |     |
| DEFINITION                | AGENCOURT_7046721 NIH_MGC_107 Homo sapiens cDNA clone IMAGE:5805819 |   |     |
| ACCESSION                 | 5', mRNA sequence.  |   |     |
| VERSION                   | BQ279131  |   |     |
| KEYWORDS                  | BQ279131.1 GI:20489339  |   |     |
| SOURCE                    | EST.  |   |     |
| ORGANISM                  | human.  |   |     |
| REFERENCE                 | Homo sapiens  |   |     |
| AUTHORS                   | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;   |   |     |
| JOURNAL                   | Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.           |   |     |
| COMMENT                   | 1 (bases 1 to 979)  |   |     |
|                           | NIH-MGC http://mgc.nci.nih.gov/.                                    |   |     |
|                           | National Institutes of Health, Mammalian Gene Collection (MGC)      |   |     |
|                           | Unpublished (1999)  |   |     |
|                           | Contact: Robert Strausberg, Ph.D.                                   |   |     |
|                           | Email: cgapbs-f@mail.nih.gov  |   |     |
|                           | Tissue Procurement: ATCC  |   |     |
|                           | cDNA Library Preparation: Rubin Laboratory                          |   |     |
|                           | cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)           |   |     |
|                           | DNA Sequencing by: Agencourt Bioscience Corporation                 |   |     |
|                           | Clone distribution: MGC clone distribution information can be       |   |     |
|                           | found through the I.M.A.G.E. Consortium/LLNL at:                    |   |     |
|                           | http://image.llnl.gov   |   |     |
|                           | Plate: L1CM2045 row: k column: 04                                   |   |     |
|                           | High quality sequence stop: 752.                                    |   |     |
| FEATURES                  | Location/Qualifiers   |   |     |
| source                    | 1..979  |   |     |
|                           | /organism="Homo sapiens"  |   |     |
|                           | /db_xref="taxon:9606"   |   |     |
|                           | /clone="IMAGE:5805819"  |   |     |
|                           | /clone_lib="NIH_MGC_107"  |   |     |
|                           | /tissue_type="adenocarcinoma, cell line"                            |   |     |
|                           | /lab_host="DH10B (phage-resistant)"                                 |   |     |
|                           | /note="Organ: breast; Vector: pORF7; Site: EcoRI;"                  |   |     |
|                           | Site 2: XhoI; cDNA made by oligo-dT priming,                        |   |     |
|                           | Directionally cloned into EcoRI/XhoI sites using the                |   |     |
|                           | following 5' adaptor: GGCAGG(G). Library constructed by             |   |     |
|                           | Ling Hong in the laboratory of Gerald W. Rubin (University          |   |     |
|                           | of California, Berkeley) using ZAP-cDNA synthesis kit               |   |     |
|                           | (Stratagene) and Superscript II RT (Life Technologies).             |   |     |
|                           | Note: This is a NIH_MGC Library."                                   |   |     |
| BASE COUNT                | 141 a 347 c 313 g 170 t   | 8 others  |     |
| ORIGIN                    |   |   |     |
| Query Match               | 40.7%; Score 258.2; DB 14; Length 979;                              |   |     |
| Best Local Similarity     | 81.5%; Pred. No. 3.5e-56;   |   |     |
| Matches 299; Conservative | 0; Mismatches 68; Indels 0; Gaps 0;                                 |   |     |
| Qy                        | 269   | CCATCTTCTGGCA@TTCCTCTCTCTGCTTCCTCAATGGTGTTGGTGCCTGAGTGGCCG 328    |     |
| Db                        | 13  | CCACCTCTGGCCTTCTCCCTCTCTGCTTCCTCTCANAGGTGGTACCCAGCTGCCCC 72       |     |
| Qy                        | 329   | GACACCTGTACCTGTCTTGGACACACCCCAGTGCCTCCACAGGGGTACCCCTGTGTCT 388    |     |
| Db                        | 73  | GACACCATGTACCTGTCCCTTCCCTCCCGATGCCCGCTGGAGTACCCCTGTGTCT 132       |     |
| Qy                        | 389   | GGATGGCTGTGGCTGTAAAGTCTGTGACGGAGGCTGGGGAGTCTCTGCCACCACT 448       |     |
| Db                        | 133   | GGATGGCTGTGGCTGTGCCGGGTATGTGCACGGCGGCTGGGGAGGCCCTTGCACCACT 192    |     |
| Qy                        | 449   | GCATGTCTGGGACCCACCGCAGGGCCTGTTTTGTACGCTTGGGCGAGGCCCTTGGCGGCCA 508 |     |

```

Db      193  CCAGCTCTGGACGCCAGCCAGGGCTGGTCTGCCAGCCGGGGAGGACCCGGTGAAG 252
Qy      509  TGGGGCTGTGTGCTCTTGGATGAGGATGAGCTAGCTGTGAGGTGAATGCCCGCAGGTA 568
Db      253  GGGGGCCCTGTGCTCTTGGCAGGAGGACAGCAGCTGTGAGGTGAACGGCCGCTGTA 312
Qy      569  CTGTGATGAGACACCTTAAACCCAAATTCGAGGTCCTCTGCGCTGTGATGACGGTGG 628
Db      313  TCGGGAGAGGGAGACCTTCCAGCCCACTGCATCATCGCTGCCGTGCGAGGACGGCG 372
Qy      629  CTTCAACC 635
Db      373  CTTCAACC 379

RESULT 14
LOCUS   BQ278961
DEFINITION BQ278961 933 bp mRNA linear EST 07-MAY-2002
5', mRNA sequence.
ACCESSION BQ278961
VERSION   BQ278961.1 GI:20489169
KEYWORDS EST.
SOURCE   human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 933)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ruben Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LCM2045 row: c column: 03
High quality sequence start: 12
High quality sequence stop: 353.
Location/Qualifiers
1. .933
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5805626"
/tissue_type="adenocarcinoma, cell line"
/notes="Organ: breast; Vector: pONB7; Site:1: EcoRI;
Site:2: XhoI; CDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GCCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
BASE COUNT 139 a 330 c 295 g 166 t
ORIGIN
Query Match 40.4%; Score 256.8; DB 14; Length 933;
Best Local Similarity 78.9%; Pred. No. 8e-56;
Matches 306; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

Qy      248  CATGAGGGGAGCCCACTGATCCATCTTGTGGCCACTTCTCTCTGCTCTCTCAAT 307
Db      16  CAAGACGAGGGACACAGCAACCTCTCTCGGCTTCTCTCTGCTCTCTCAAA 75
Qy      308  GGTGTGTGCCAGCTGTGCGGAGACACCTGCTACTGCTGTGGACACCAACCCAGTGC 367
Db      76  GGTGCGTACCAGCTGTGTCCGACACCATGATGCTGCCCTGCGCCACTCCCGATGCC 135

```

```

Qy      368  ACAGGGGGTACCCCTGGTGTGATGGCTGTGTAAAGTGTGTGACGAGGCT 427
Db      136  GCTGGGAGTACCCCTGGTGTGATGGCTGTGTGCTGTGCGGGTATGTGCACGGGCT 195
Qy      428  GGGGGAGTCTTGGACACACTGATGTGTGCGACCCAGCCAGGGGCTGGTTGTGACGC 487
Db      196  GGGGGAGCCTTGGGACCAATCCACGTCTGCGACGCGCAGCGGGCTGGTCTGCCAGCC 255
Qy      488  TGGGGCAGGCGCTTGGCGGCATGGGCTGTGTCTCTTGGATGAGGATGACGGTAGCTG 547
Db      256  CGGGGAGGACCGCGGTGGAGGGGGCCCTGTGCTCTGTGTGTGACGAGCAGACGAGCTG 315
Qy      548  TGAGGTGAATGGCCGAGGTACCTGGATGAGAGACCTTTAAACCAATTTGCAGGGTCT 607
Db      316  TGAGGTGAACGGCGCGCTGTATCGGAAGGGAGAGACCTTTCAGCCCACTGCAGCATCCG 375
Qy      608  GTCCCGCTGTGATGACGGTGGCTTCAAC 635
Db      376  CTCCCGCTGCGAGGACGGCGGCTTCAAC 403

RESULT 15
LOCUS   BQ073722
DEFINITION BQ073722 1073 bp mRNA linear EST 02-APR-2002
5', mRNA sequence.
ACCESSION BQ073722
VERSION   BQ073722.1 GI:19902768
KEYWORDS EST.
SOURCE   human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1073)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ruben Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LCM2047 row: k column: 19
High quality sequence stop: 517.
Location/Qualifiers
1. .1073
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5806602"
/tissue_type="epidermoid carcinoma, cell line"
/notes="Organ: lung; Vector: pONB7; Site:1: EcoRI; Site:2:
XhoI; CDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."
BASE COUNT 166 a 384 c 347 g 176 t
ORIGIN
Query Match 40.3%; Score 256; DB 14; Length 1073;
Best Local Similarity 81.9%; Pred. No. 1.4e-55;
Matches 295; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

Qy      276  CTGGCCACTTCCTTCTCTCTCTCTCTCAATGGTGTGCCAGCTGTGCCGACACC 335

```

|     |  |     |
|-----|--|-----|
| 2   | CTGGGCTTCTCCCTCCCTCTCTGCTCTCTCTCAAAGGTCGTAACCGAGCTGTGCCGACACCA | 61  |
| 336 | TGTACTCTCTTGGACACACCCACAGTCGCCACAGGGGTACCCCTGGTCTGTGATGAC      | 395 |
| 62  | TGTACTGCCCTTGCCACACTCCCGATGCCGCTGGAGTACCCCTGGTCTGTGATGAC       | 121 |
| 336 | TGTGCTCTCTAAAGTGTGCACGAGGCTGGGGAGTCTCTGGACCACTCGATGTC          | 455 |
| 122 | TGTGCTCTGCCGGGTATGTGACGGCGCTGGGGAGCCCTGCGACCACTCCACGTC         | 181 |
| 456 | TGCGACCCAGCCAGGCGCTGGTTGTCAGCTGGGCGAGGCCCTGGCGGCCATGGGCT       | 515 |
| 182 | TGCGAGCCAGCCAGGCGCTGGTCTGCCAGCCCGGGCAGGACCGGTGGCCGGGGGGCC      | 241 |
| 516 | CTGTGTCTCTTGGATGAGGATGACGCTAGCTGTGAGGTGAATGGCCGCGCAGGTACCTGGAT | 575 |
| 242 | CTGTGCTCTTGGCAGAGGACGACAGCAGCTGTGAGTGAACGGCGCGCTGTATCGGNA      | 301 |
| 576 | GGAGAGACCTTTAAACCAATTGACGGTCCCTGTGGCTGTGTGATGACGGTGGCTTAC      | 635 |
| 302 | GGGAGAGCTTCCAGCCCCACTCAGCATCCGCTGGCGTGGAGGACGGCGGCTTAC         | 361 |

Search completed: July 29, 2003, 02:51:28  
Job time : 1128.2 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 15:59:01 ; Search time 2454.08 Seconds  
(without alignments)  
10471.464 Million cell updates/sec

Title: US-10-010-408-1\_COPY\_1\_883  
Perfect score: 883  
Sequence: 1 GAGCGTCTGATCTCCAGAG.....ACCACCTGTGGCTGGGCAT 883

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_em.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_om.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pl.\*
- 26: em\_ro.\*
- 27: em\_sts.\*
- 28: em\_un.\*
- 29: em\_vi.\*
- 30: em\_htg\_hum.\*
- 31: em\_htg\_inv.\*
- 32: em\_htg\_other.\*
- 33: em\_htg\_mus.\*
- 34: em\_htg\_pln.\*
- 35: em\_htg\_rod.\*
- 36: em\_htg\_mam.\*
- 37: em\_htg\_vrt.\*
- 38: em\_sy.\*
- 39: em\_htgo\_hum.\*
- 40: em\_htgo\_mus.\*
- 41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID        | Description        |
|------------|-------|-------------|--------|----|-----------|--------------------|
| 1          | 821.4 | 93.0        | 1741   | 10 | AF259981  | AF259981 Rattus no |
| 2          | 727.6 | 82.4        | 1734   | 6  | AR210324  | AR210324 Sequence  |
| 3          | 727.6 | 82.4        | 1734   | 6  | AR210325  | AR210325 Sequence  |
| 4          | 727.6 | 82.4        | 1734   | 10 | AF100778  | AF100778 Mus muscu |
| 5          | 718.2 | 81.3        | 1739   | 10 | AF126063  | AF126063 Mus muscu |
| 6          | 434.6 | 49.2        | 1266   | 6  | AX076919  | AX076919 Sequence  |
| 7          | 434.6 | 49.2        | 1266   | 6  | AX464186  | AX464186 Sequence  |
| 8          | 434.6 | 49.2        | 1283   | 9  | AF083500  | AF083500 Homo sapi |
| 9          | 434.6 | 49.2        | 1293   | 6  | AR210322  | AR210322 Sequence  |
| 10         | 434.6 | 49.2        | 1293   | 6  | AR210323  | AR210323 Sequence  |
| 11         | 434.6 | 49.2        | 1309   | 9  | AF074604  | AF074604 Homo sapi |
| 12         | 434.6 | 49.2        | 1427   | 9  | AF100780  | AF100780 Homo sapi |
| 13         | 434.6 | 49.2        | 1450   | 9  | BC017782  | BC017782 Homo sapi |
| 14         | 423   | 47.9        | 738    | 6  | AR210337  | AR210337 Sequence  |
| 15         | 420.6 | 47.6        | 841    | 6  | AR210338  | AR210338 Sequence  |
| 16         | 256.6 | 29.1        | 137964 | 2  | AC126895  | AC126895 Rattus no |
| 17         | 221.6 | 25.1        | 137964 | 2  | AC126895  | AC126895 Rattus no |
| 18         | 221.6 | 25.1        | 226303 | 2  | AC095418  | AC095418 Rattus no |
| 19         | 216.2 | 24.5        | 61072  | 10 | AL731698  | AL731698 Mouse DNA |
| 20         | 199.2 | 22.6        | 216757 | 2  | AL669906  | AL669906 Mus muscu |
| 21         | 169.8 | 19.2        | 107260 | 9  | AL139352  | AL139352 Human DNA |
| 22         | 138.8 | 15.7        | 1805   | 5  | CHKCEF    | J04496 Chicken CEF |
| 23         | 128.6 | 14.6        | 1585   | 5  | NVI271167 | J04496 Chicken CEF |
| 24         | 128.4 | 14.5        | 2075   | 6  | AR018957  | AR018957 Notophtha |
| 25         | 128.4 | 14.5        | 2075   | 6  | AR118595  | AR118595 Sequence  |
| 26         | 128.4 | 14.5        | 2075   | 6  | AR130891  | AR130891 Sequence  |
| 27         | 128.4 | 14.5        | 2075   | 6  | AX206708  | AX206708 Sequence  |
| 28         | 128.4 | 14.5        | 2075   | 6  | I11636    | I11636 Sequence 1  |
| 29         | 128.4 | 14.5        | 2075   | 6  | I32210    | I32210 Sequence 1  |
| 30         | 128.4 | 14.5        | 2075   | 9  | HUMCONGRO | M92934 Homo sapien |
| 31         | 128.4 | 14.5        | 2312   | 9  | HSCCTGF   | X78947 H.sapiens m |
| 32         | 128.4 | 14.5        | 2998   | 6  | AR119211  | AR119211 Sequence  |
| 33         | 128.4 | 14.5        | 2998   | 6  | AR151276  | AR151276 Sequence  |
| 34         | 127.4 | 14.4        | 1975   | 6  | A28405    | A28405 Sequence I  |
| 35         | 127.4 | 14.4        | 1975   | 6  | A28444    | A28444 nov mRNA se |
| 36         | 127.4 | 14.4        | 1976   | 5  | GGNOVMRNA | X59284 G.gallus no |
| 37         | 126.6 | 14.3        | 2052   | 9  | HSCYR61   | Y11307 H.sapiens C |
| 38         | 125.8 | 14.2        | 2338   | 6  | E37595    | E37595 Monoclonal  |
| 39         | 125.8 | 14.2        | 2338   | 10 | AB023068  | AB023068 Rattus no |
| 40         | 125.8 | 14.2        | 2345   | 6  | AX402019  | AX402019 Sequence  |
| 41         | 125.8 | 14.2        | 2345   | 10 | AF120275  | AF120275 Rattus no |
| 42         | 125.8 | 14.2        | 2350   | 6  | AR194011  | AR194011 Sequence  |
| 43         | 125   | 14.2        | 1146   | 6  | E13814    | E13814 cDNA encodi |
| 44         | 125   | 14.2        | 1887   | 9  | AF031385  | AF031385 Homo sapi |
| 45         | 125   | 14.2        | 1935   | 9  | AF003594  | AF003594 Homo sapi |

ALIGNMENTS

|            |  |            |      |        |                 |
|------------|--|------------|------|--------|-----------------|
| RESULT 1   | AF259981   | 1741 bp    | mRNA | linear | ROD 09-MAY-2000 |
| LOCUS      | Rattus norvegicus CCN family protein COP-1 (Cop-1) mRNA, complete cds. |            |      |        |                 |
| DEFINITION | AF259981   |            |      |        |                 |
| ACCESSION  | AF259981   |            |      |        |                 |
| VERSION    | AF259981.1   | GI:7739780 |      |        |                 |
| KEYWORDS   |  |            |      |        |                 |
| SOURCE     | Rattus norvegicus.   |            |      |        |                 |
| ORGANISM   | Rattus norvegicus  |            |      |        |                 |
|            | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;      |            |      |        |                 |
|            | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;          |            |      |        |                 |
|            | Rattus.  |            |      |        |                 |
| REFERENCE  | 1 (bases 1 to 1741)  |            |      |        |                 |
| AUTHORS    | Zhang,R., Averboukh,L., Zhu,W., Zhang,H., Jo,H., Dempsey,P.J.,         |            |      |        |                 |

Coffey R.J., Pardee A.B. and Liang, P.  
Identification of rCop-1, a new member of the CCN protein family,  
as a negative regulator for cell transformation  
Mol. Cell. Biol. 18 (10), 6131-6141 (1998)  
98414629  
PUBMED  
9742130  
REFERENCE  
2 (bases 1 to 1741)  
Liang, P.  
Direct Submission  
Submitted (24-APR-2000) Cell Biology, Vanderbilt-Ingram Cancer  
Center, 649 MRB II, Nashville, TN 37232, USA  
FEATURES  
source  
1..1741  
/organism="Rattus norvegicus"  
/db\_xref="taxon:10116"  
1..1741  
/gene="Cop-1"  
262..1014  
/gene="Cop-1"  
/note="secreted protein"  
/codon\_start=1  
/product="CCN family protein COP-1"  
/protein\_id="AAF6901.1"  
/db\_xref="GI:7739781"  
/translation="MRGSLIRLILATSLFLLSWCAQLRTPCTPWTTPQCPGV  
LVDCGCGCKVCARLHVCEPQGLVCPQGGHGVCLLDDGDCV  
NGRRYLDGETFKNRVLCDGDFGLPLCSEDTPLPSWDPRPKRIQVPGKCP  
WYCDGVPTAORSAAOGLSALVTPASADAPWNSWANGPCSTGLGLIATRVN  
QNRFCLEIQRLLCPRLAARSHSNNAF"  
BASE COUNT 386 a 491 c 480 g 384 t  
ORIGIN  
Query Match 93.0%; Score 821.4; DB 10; Length 1741;  
Best Local Similarity 97.8%; Pred. No. 2,1e-188;  
Matches 864; Conservative 0; Mismatches 16; Indels 3; Gaps 3;  
QY 1 GACGCTTCGTACTCCAGAGACCCTGGGGTGGGACAGGGCTTGGCAAGGCTGCAGCC 60  
DB 17 GACGCTTCGTACTCCAGAGACCCTGGGGTGGGACAGGGCTTGGCAAGGCTGCAG 74  
QY 61 GCTGGGAGTGGCTGGGAATGAGGCTTTATTACTGGGAAGTGGAGGCTTAAGAGGCTC 120  
DB 75 GCTGGGAGTGGCTGGGAATGAGGCTTTATTACTGGGAAGTGGAGGCTTAAGAGGCTC 134  
QY 121 CTGTCAGCTTGTCCTAAAGTCTTAGCAGCTTGTTGGTGGCTTTCACACTGTCAGA 180  
DB 135 CTGTCAGCTTGTCCTAAAGTCTTAGCAGCTTGTTGGTGGCTTTCACACTGTCAGA 194  
QY 181 CACCTTCGTGGTGGCTCCAGGGCTCAGCTTCAGCTTGAAGCTGGCTCCACAGGGAC 240  
DB 195 CACCTTCGTGGTGGCTCCAGGGCTCAGCTTCAGCTTGAAGCTGGCTCCACAGGGAC 253  
QY 241 ACGGTGACATGAGGGGACGCCACCTGATCCATCTTCTGGCCACTTCTCTCTCCCTTC 300  
DB 254 ACGGTGACATGAGGGGACGCCACCTGATCCGCTTCTGGCCACTTCTCTCTCTCCCTTC 313  
QY 301 TCTCAATGTTGTGCCAGCTGTGCCGACACCCCTGTACCTGTCTTGGACACACCC 360  
DB 314 TCTCAATGTTGTGCCAGCTGTGCCGACACCCCTGTACCTGTCTTGGACACACCC 373  
QY 361 AGTGGCCACAGGGGTACCCCTGGTGGTGGATGGCTGGCTGTAAAGTGTGGAC 420  
DB 374 AGTGGCCACAGGGGTACCCCTGGTGGTGGATGGCTGGCTGTAAAGTGTGGAC 433  
QY 421 GGAGGCTGGGAGTCCTGGACACCTGCATGTCTGGACGCCACCCAGGGGCTGGTT 480  
DB 434 GGAGGCTGACGAGTCTTGGAAACCTGCATGTCTGGACGCCACCCAGGGGCTGGTT 493  
QY 481 CTCAGCCTGGGACGGCCCTGGCGCCATGGGGCTGTGTCTCTTGGATGAGGATGACG 540  
DB 494 CTCAGCCTGGGACGGCCCTGGCGCCATGGGGCTGTGTCTCTTGGATGAGGATGACG 553  
QY 541 GTAGCTGTGAGGTGAATGGCGCAGGTACCTGGATGGAGACCTTTAAACCCAAATTGCA 600

Db 554 CTGACTGTGAGTGAATGGCGCAGGTACCTGGATGGAGAGACCTTTAAACCCAAATTGCA 613  
QY 601 GGGTCTCTGGCGCTGTGATGACGCTGGCTTCCACCTGCTGCGCTGTGCACTGAGGATG 660  
Db 614 GGGTCTCTGGCGCTGTGATGACGCTGGCTTCCACCTGCTGCGCTGTGCACTGAGGATG 673  
QY 661 TCGGCTGCCAGCTGGGACTGCCACGCCGCCAAGAGAAATACAGGTGCCAGGAAAGTCT 720  
Db 674 TGACGCTGCCAGCTGGGACTGCCACGCCGCCAAGAGAAATACAGGTGCCAGGAAAGTCT 733  
QY 721 GCCCGAGTGGTATGTGACAGGAGTGCACACCGGCGATCCAGGCTCCAGGCGCGCAAG 780  
Db 734 GCCCGAGTGGTATGTGACAGGAGTGCACACCGGCGATCCAGGCTCCAGGCGCGCAAG 793  
QY 781 GACACCAACTTCTGCCCTTGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840  
Db 794 GACACCAACTTCTGCCCTTGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 853  
QY 841 GCACAGCCTGGGGCCCTGCTCAACACACTGTGGGCTGGGCAT 883  
Db 854 GCACAGCCTGGGGCCCTGCTCAACACACTGTGGGCTGGGCAT 896  
RESULT 2  
AR210324  
LOCUS 1734 bp DNA linear PAT 20-JUN-2002  
DEFINITION Sequence 17 from patent US 6387657.  
ACCESSION AR210324  
VERSION AR210324.1 GI:21512525  
KEYWORDS Unknown.  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 1734)  
AUTHORS Bokstein, D.A., Cohen, R.L., Goddard, A.D., Gurney, A.L., Hillan, K.J.,  
Lawrence, D.A., Levine, A.J., Pennica, D., Roy, M. Ann. and Wood, W.I.  
TITLE WISP polypeptides and nucleic acids encoding same  
JOURNAL Patent: US 6387657-A 17 14-MAY-2002;  
FEATURES Location/Qualifiers  
source 1..1734  
BASE COUNT 355 a 491 c 495 g 393 t  
ORIGIN  
Query Match 82.4%; Score 727.6; DB 6; Length 1734;  
Best Local Similarity 91.9%; Pred. No. 1,1e-165;  
Matches 816; Conservative 0; Mismatches 59; Indels 13; Gaps 4;  
QY 3 CGCTTCTGATCTCCAGAGGACCCCTGGGTGGGACAGGGCCCTTGGCAAGGCTGCAGCGCC 62  
DB 13 CGCTTCTGATCTCCAGAGGACCCCGGCTGGGACAGGGCCCTTGGCGAGGCTGCAGCTGC 72  
QY 63 TG-GCAGTGGCTTGGGAATGAGGCTTTATTACTGGGAAGTGGAGGCTTAAGAGGCTCC 121  
DB 73 TGTGGCAGTAGCTGGGATGGAGTCTTTCTTGTGGGAAGTGGAGGCTGAGAGGCTCC 132  
QY 122 TGTGAG---CTTGTCTCTAAAGTCTTAGCAGCTTGTGGTGGCTTGGGCTTTCACACACTGTCA 178  
DB 133 TGTGAGGCTCTCTGCTTAACTCTTGGCAGCTTGGGCTGGCTTGGGCTTTCACACACTGTCA 192  
QY 179 GACACCTTCTGGTGGCTCCAGGCGCTCACCTTTCAGTCTTGAAGCTGGCTCCACAGGG 238  
DB 193 GACACCTTCTGGTGGCTCTCGGCC-----TCAGGTTGAAGCTGGCTCCACAGGG 246  
QY 239 ACAGGTCACATGAGGGGCGACCCACTGATCCATCTTCTGGCAGCTTCTTCTCTGCT 298  
DB 247 ACAGGTCACATGAGGGGCGACCCACTGATCATCTTCTGGCCATTTCTTCTCTGCT 306  
QY 299 TCTCTCAATGTGTGCGCAGCTGTGCGGACACCCCTGTACCTGCTCTCTTGGACACACC 358  
DB 307 TCTCTCAATGTGTATTCCAGCTGTGCGGACACCCCTGTGCTCTGCTCTTGGACACACC 366

```
QY 359 CCAAGTGGCCACAGGGGGTACCCCTGGGTGCTGGATGGCTGTGGCTGTGAAGTGTGTGC 418
Db 367 CCAAGTGGCCACAGGGGGTACCCCTGGGTGCTGGATGGCTGTGGCTGTGAAGTGTGTGC 426
QY 419 ACGGAGGCTGGGAGTCTCGACACCACTGCATGCTGTGGACCCAGCCAGGCGCTGGT 478
Db 427 ACGGAGGCTGGGAGTCTCGACACCACTGCATGCTGTGGACCCAGCCAGGCGCTGGT 486
QY 479 TTGTACAGCTGGGACAGCCCTGGCGGCTTACCTGCTGTGGATGAGGATGA 538
Db 487 TTGTACAGCTGGGACAGCCCACTGGCGGCTTACCTGCTGTGGATGAGGATGA 546
QY 539 CGGTAGCTGTGAGTGAATGCGCGAGTACTCGATGGAGAGACCTTTAAACCCAATTG 598
Db 547 CGGAGGCTGTGAGTGAATGCGCGAGTACTCGATGGAGAGACCTTTAAACCCAATTG 606
QY 599 CAGGCTGTGGCGCTGTGATGACGGTGGCTTACCTGCTGGCTGCTGGCTGTGCAGTGAGGA 658
Db 607 CAGGCTGTGGCGCTGTGATGACGGTGGCTTACCTGCTGGCTGCTGGCTGTGCAGTGAGGA 666
QY 659 TGTGGGCTGCCAGCTGGGACTGCCACGCCCCCAAGAGATACAGGTGCGCAGGAAAGTG 718
Db 667 TGTGGGCTGCCAGCTGGGACTGCCACGCCCCCAAGAGATACAGGTGCGCAGGAAAGTG 726
QY 719 CTGCCCCAGTGGTGTGACACAGGAGTGA---CACCGCGATCCAGCGCTCCACGGC 775
Db 727 CTGCCCCAGTGGTGTGACACAGGAGTGA---CACCGCGATCCAGCGCTCCACGGC 786
QY 776 GCAAGGACACCAACTTTCTGCCCTGTACACCTCCCTGCTGCTGATGCTGCTGCTCCAAA 835
Db 787 CCAAGGACACCAACTTTCTGCCCTGTACACCTCCCTGCTGCTGATGCTGCTGCTCCAAA 846
QY 836 TTGGAGCACAGCTGGGCGCCCTGCTCAACCACTGTGGGTGGGCGAT 883
Db 847 CTGGAGCACAGCTGGGCGCCCTGCTCAACCACTGTGGGTGGGCGAT 894

RESULT 3
AR210325/c AR210325 1734 bp DNA linear PAT 20-JUN-2002
LOCUS Sequence 18 from patent US 6387657.
ACCESSION AR210325
VERSION AR210325.1 GI:21512526
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
AUTHORS 1 (bases 1 to 1734)
Botstein, D.A., Cohen, R.L., Goddard, A.D., Gurney, A.L., Hillan, K.J.,
Lawrence, D.A., Levine, A.J., Pennica, D., Roy, M., Ann. and Wood, W.I.
TITLE WISP polypeptides and nucleic acids encoding same
JOURNAL Patent: US 6387657-A 18 14-MAY-2002;
FEATURES
source Location/Qualifiers
BASE COUNT 393 a 495 c 491 g 355 t
ORIGIN
Query Match 82.4%; Score 727.6; DB 6; Length 1734;
Best Local Similarity 91.9%; Pred. No. 1.1e-165;
Matches 816; Conservative 0; Mismatches 59; Indels 13; Gaps 4;

QY 3 CGCTTCTGATCTCCAGAGACCCCTGGGTGGGACAGGGGCGCTTGGCAAGGCTGCAGCGC 62
Db 1722 CGCTTCTGATCTCCAGAGACCCCGGGCTGGGACAGGGGCGCTTGGGAGGCTGCAGTGC 1663
QY 63 TG-GGCACTGGCTGGAAATGAGGTCTTTATTACTGGGAATGAGGAGCTTAAGAGCTCC 121
Db 1662 TGTGCACTAGCTGGGATGAGGTCTTTCTTGTGGAACTGAGGAGCTGAGAGGCTCC 1603
QY 122 TGTACAG---CTTGCTCTAAAGCTTAGCATTGCTGGCTTGGCTTCACACACTGTCA 178
Db 1602 TGTACAGCTCTGTCTTAAAGCTTGGCACTTGGGCTTGGGCTTTCACACACTGTCA 1543
```

```
QY 179 GACACCTTCGTGGTGGCTCCACGGGCTCACCTTCAGTTTGAAGTGGCTCCACAAGG 238
Db 1542 GACACCTTCGTGGTGGCTCCTCGGCC-----TCAGTTTGAAGTGGCTCCACAAGG 1489
QY 239 ACAGGTGACATGAGGGGAGCCCACTGATCCATCTTCTGGCCACTTCCTTCTCTGCCT 298
Db 1488 ACAGGTGACATGAGGGGAGCCCACTGATCCATCTTCTGGCCACTTCCTTCTCTGCAT 1429
QY 299 TCTCTCAATGGTGTGCCAGCTGTGCCGACACCTGTACCTGTGCTCTTGGACACACCC 358
Db 1428 TCTCTCAATGGTGTATTCACAGCTGTGCCAGACACCTGTGCTCTTGGACACACCC 1369
QY 359 CCAAGTGGCCACAGGGGTACCCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 418
Db 1368 CCAAGTGGCCACAGGGGTACCCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1309
QY 419 ACGGAGCTGGGAGTGTCTGGACACCTGCATGCTGTGGACCCAGCCAGGCGCTGGT 478
Db 1308 ACGGAGCTGGGAGTGTCTGGACACCTGCATGCTGTGGACCCAGCCAGGCGCTGGT 1249
QY 479 TTGTACAGCTGGGACAGCCCTGGCGGCTTACCTGCTGTGGTGGTGGTGGTGGTGG 538
Db 1248 TTGTACAGCTGGGACAGCCCGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1189
QY 539 CGGTAGCTGTGAGTGAATGCGCGAGTACTCGATGGATGGAGAGACCTTTAAACCCAATTG 598
Db 1188 CGGAGCTGTGAGTGAATGCGCGAGTACTCGATGGATGGGAGACCTTTAAACCCAATTG 1129
QY 599 CAGGCTGTGCGGCTGTGATGACGGTGGCTTACCTGCTGCTGCTGCTGCTGCTGCTG 658
Db 1128 CAGGCTGTGCGGCTGTGATGACGGTGGCTTACCTGCTGCTGCTGCTGCTGCTGCTG 1069
QY 659 TGTGCGGCTGCCAGCTGGGACTGCCACGCCCCCAAGAGATACAGGTGCGCAGGAAAGTG 718
Db 1068 TGTGCGGCTGCCAGCTGGGACTGCCACGCCCCCAAGAGATACAGGTGCGCAGGAAAGTG 1009
QY 719 CTGCCCCAGTGGGTATGTGACACAGGAGTGA---CACCGCGATCCAGCGCTCCACGGC 775
Db 1008 CTGCCCCAGTGGGTATGTGACACAGGAGTGA---CACCGCGCAATCCAGCGCTCCACGC 949
QY 776 GCAAGGACACCAACTTTCTGCCCTGTGACCTCCCTGCTGCTGCTGCTGCTGCTGCTG 835
Db 948 CCAAGGACACCAACTTTCTGCCCTGTGACCTCCCTGCTGCTGCTGCTGCTGCTGCTG 889
QY 836 TTGGAGCACAGCTGGGCGCCCTGCTCAACCACTGTGGGTGGGCGAT 883
Db 888 CTGGAGCACAGCTGGGCGCCCTGCTCAACCACTGTGGGTGGGCGAT 841

RESULT 4
AF100778 1734 bp mRNA linear ROD 17-DEC-1998
LOCUS Mus musculus connective tissue growth factor related protein WISP-2
DEFINITION (Wisp2) mRNA, complete cds.
ACCESSION AF100778
VERSION AF100778.1 GI:4028578
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus
REFERENCE
AUTHORS 1 (bases 1 to 1734)
Pennica, D., Swanson, T.A., Welsh, J.W., Roy, M.A., Lawrence, D.A.,
Lee, J., Brush, J., Taneyhill, L.A., Deuel, B., Lew, M., Watanabe, C.,
Cohen, R.L., Melnick, M.F., Finley, G.G., Quirke, P., Goddard, A.D.,
Hillan, K.J., Gurney, A.L., Botstein, D. and Levine, A.J.
WISP genes are members of the connective tissue growth factor
family that are up-regulated in wnt-1-transformed cells and
aberrantly expressed in human colon tumors
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 95 (25), 14717-14722 (1998)
MEDLINE 99061933
PUBMED 9843955
```

REFERENCE 2 (bases 1 to 1734)  
AUTHORS Pennica, D.  
TITLE Direct Submission  
JOURNAL Submitted (23-OCT-1998) Molecular Oncology, Genentech, Inc., 1 DNA Way, South San Francisco, CA 94080, USA  
FEATURES  
source Location/Qualifiers  
1..1734  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/cell\_line="C57MG"  
/cell\_type="epithelial"  
/tissue\_type="mammary"  
/note="transformed by Wnt-1"  
1..1734  
/gene="Wisp2"  
257..1012  
/gene="Wisp2"  
/codon\_start=1  
/product="connective tissue growth factor related protein WISP-2"  
/protein\_id="AAC96320.1"  
/db\_xref="GI:4028579"  
/translation="MRGNPLIHLAIFLCLILSMVYSLCPAPCAPWTPPQCPGVP  
LYVDCGCRVCRARLGVSCDHLHVCDFQGLVCPGAGPSGRGAVCLFEDDSCEV  
NGRRYLDGETFKPNCRVLCRDGDFCLPLCSEDVRLPSWDPRPRIOVPGRCPE  
WYCDQAVQPAIQPSSAQHLSALVTPASADGPCPNWSTAWGPCSTTCGLGIATRV  
NONRFOCLEIQRLLCLSRPLASRSHGWSNAF"  
BASE COUNT 355 a 491 c 495 g 393 t  
ORIGIN  
Query Match 82.4%; Score 727.6; DB 10; Length 1734;  
Best Local Similarity 91.9%; Pred. No. 1.1e-165;  
Matches 816; Conservative 0; Mismatches 59; Indels 13; Gaps 4;  
QY 3 CCCTTCATCTCCAGAGGACCCCTGGGTGGGACAGGGCCCTTGGCAAGGCTGCAGCCGC 62  
DB 13 CGCTCTGATCTCCAGAGGACCCCGGGTGGGACAGGGCCCTTGGCGAGGCTGCAGCTGC 72  
QY 63 TG-GGCAGTGGCTTGGAAATGAGGCTTTATTACTGGGAACCTAGAGAGCTAAGAGCTCC 121  
DB 73 TGTGCGAGTAGCTGGGATGAGGCTTTCTGTGTGGGAACCTAGAGAGCTGCAGGCTCC 132  
QY 122 TGTCAG---CTTGTCTTAAAGTCTTAGCACTTGTGGTGGCTTGGGCTTCACACACTGTCA 178  
DB 133 TGTCAAGCTCTGTCTTAACTCTTGGCACTTGGCGTGGCTTGGGCTTCACACACTGTCA 192  
QY 179 GACACCTTCGTGGTGGCTCCACAGGCTCACCTTCAGGTTGAAGCTGCTCCACCAAGG 238  
DB 193 GACACCTTCGTGGTGGCTCCCTCGGCC-----TCAGTTTGAAGCTGGCTCCACCAAGG 246  
QY 239 ACACGGTGACATGAGGGGACGCCACTGATCATCTTGGCCACTTCTTCTCTCTGCT 298  
DB 247 ACACGGTGACATGAGGGGACGCCACTGATCATCTTGGCCACTTCTTCTCTCTGAT 306  
QY 299 TCTCTCAATGTTGTGGCGAGCTTGGCGGACACCCCTGTACCTGCTCTCTTGGACACACC 358  
DB 307 TCTCTCAATGTTGTATCCAGCTGTGCCACACCCCTGTGCTCTCTTGGACACACC 366  
QY 359 CCAGTGCCACAGGGGGTACCCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 418  
DB 367 CCAGTGCCACAGGGGGTACCCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 426  
QY 419 ACGGAGGTGGGGGAGTCTTGGACACCTGATGCTGTGCGACCCACCCAGCCAGGGGCTGGT 478  
DB 427 ACGGAGGTGGGGGAGTCTTGGACACCTGATGCTGTGCGACCCACCCAGCCAGGGGCTGGT 486  
QY 479 TTGTCAGCTGGGGGAGGCTTGGCGGCGCATGGGGCTGTGTCTCTCTGATGAGGATGA 538  
DB 487 TTGTCAGCTGGGGGAGGCTTGGCGGCGCATGGGGCTGTGTCTCTCTGATGAGGATGA 546  
QY 539 CGGTAGCTGTGAGGTGAATGGCCGCGAGGTACCTGATGGAGAGACCTTTAAACCCAAATTG 598  
DB 547 CGGGAGCTGTGAGGTGAATGGCCGCGAGGTACCTGATGGGAGAGACCTTTAAACCCAAATTG 606

QY 599 CAGGGTCTCTGTCGCCCTGTGATGAGGTGGCTTACCTGCCCTGCGCTGTGCAGTGAGGA 658  
DB 607 CAGGGTTTGTGCCCTGTGATGAGGTGGCTTACCTGCCCTGCGCTGTGCAGTGAGGA 666  
QY 659 TGTGGGCTGCGCTAGCTGGGACTGCCAGGCCCAAGAGAAATACAGGTGCCAGGAAGTG 718  
DB 667 TGTGGGCTGCGCTAGCTGGGACTGCCAGGCCCAAGAGAAATACAGGTGCCAGGAAGTG 726  
QY 719 CTGCCCCGAGTGGTATGTGACCCAGGGAGTGA---CACGGGATCCAGCGCTCCACGGC 775  
DB 727 CTGCCCCGAGTGGTATGTGACCCAGGGAGTGA---CACGGGATCCAGCGCTCCCTCAGC 786  
QY 776 GCAAGGACACCAACCTTCTGCGCTTGTCACTCTCTCTCTCTCTCTCTCTCTCTCTCT 835  
DB 787 CCAAGGACACCAACCTTCTGCGCTTGTCACTCTCTCTCTCTCTCTCTCTCTCTCTCT 846  
QY 836 TTGGAGCACACAGCTGGGGCCCTGTCTCAACCACTGTGGGTGGGCAT 883  
DB 847 CTGGAGCACAGCTGGGGCCCTGTCTCAACCACTGTGGGTGGGCAT 894  
RESULT 5  
AF126063 1739 bp mRNA linear ROD 12-OCT-1999  
LOCUS Mus musculus connective tissue growth factor-like protein precursor  
DEFINITION (Ctgef) mRNA, complete cds.  
ACCESSION AF126063  
VERSION AF126063.1 GI:4337059  
KEYWORDS Mus musculus.  
SOURCE Mus musculus  
ORGANISM Mus musculus  
REFERENCE 1 (bases 1 to 1739)  
AUTHORS Kumar, S., Hand, A.T., Connor, J.R., Dodds, R.A., Ryan, P.J., Trill, J.J., Fisher, S.M., Nuttall, M.E., Lipshutz, D.B., Zou, C., Hwang, S.M., Voets, B.J., James, I.E., Rieman, D.J., Gowen, M. and Lee, J.C.  
TITLE Identification and cloning of a connective tissue growth factor-like cDNA from human osteoblasts encoding a novel regulator of osteoblast functions  
JOURNAL J. Biol. Chem. 274 (24), 17123-17131 (1999)  
MEDLINE 99287915  
PUBMED 10358067  
REFERENCE 2 (bases 1 to 1739)  
AUTHORS Kumar, S. and Zou, C.  
TITLE Direct Submission  
JOURNAL Submitted (04-FEB-1999) Bone & Cartilage Biology, UW 2109, SmithKline Beecham, 709 Swedeland Rd., King of Prussia, PA 19406, USA  
FEATURES  
Location/Qualifiers  
1..1739  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/tissue\_type="lung"  
1..1739  
/gene="Ctgef"  
242..997  
/gene="Ctgef"  
/note="similar to the Mus musculus WISP-2 protein encoded by the sequence presented in GenBank Accession Number AF100778; putative growth factor; CTGF-L; contains IGF binding (IGFBP), Von Willebrand Factor type C (VWC) repeat and thrombospondin type I (TSPI) domains; member of the CCN (CtGF/Cyrf61/Nov) family; lacks the fourth carboxy-terminal (CT) domain present in other members of the CCN family"  
/codon\_start=1  
/product="connective tissue growth factor-like protein precursor"  
protein\_id="AADI8058.1"  
/db\_xref="GI:4337060"



/translation="MRGNBLIHLAISFLICILSMVYSQICLPAPCAPWTPPQCPGPV  
LVLDDGCRVRCRDLHVCDFQSGVLPQPGAGPGRGAVCLFEEDDGSCEV  
NGRRLDGGTFKPNCRVLCRDGDFCLPLGSEDVRLPSWDCPRPRRTQVGRCCPE  
WVCDQAVMOPATOPSSAOGHQLSALVTPASADGPCPNWSTWANGPCSTTCCGLGIATRV  
NQNRFCLEIQRRLCLRSRPLASRSHGWSNAF"

BASE COUNT 375 a 480 c 489 g 395 t  
ORIGIN

Query Match 81.3%; Score 718.2; DB 10; Length 1739;  
Best Local Similarity 91.4%; Pred. No. 2.1e-163;  
Matches 809; Conservative 0; Mismatches 63; Indels 13; Gaps 4;

QY 6 TTCTGATCTCCAGAGACCCCTGGGTGGGACAGAGGGCCCTTGGCAAGCTGTCAGCCGCTG- 64  
Db 1 TCCTGATCTCCAGAGACCCCGGCTGGGACAGGGCCCTTGGGAGGCTGTCAGCTGCT 60  
QY 65 GGCAGTGGCTTGAATGGAGTCTTTATTACTGGGAACCTGAGGAGCTAAGAGGCTCCTGT 124  
Db 61 GGCAGTAGCTTGGGATGGAGTCTTTCTTCTGGGAACCTGAGGAGCTGAGAAGCTCCTGT 120  
QY 125 CAG---CTTGCTCTAAAGTCTTAGCACTTTGTGTGGCTTGGGCTTCCACACACTGTGCAGAC 181  
Db 121 CAGGCTCCTGCTCTAAACTCTTGGCACTTGGCGTGGCTTGGGCTTCCACACACTGTGCAGAC 180  
QY 182 ACCTTGGTGGTCCAGGCTCCAGGCTCCTACCTTCAAGTTTGAAGCTGGCTCCACAAGGACA 241  
Db 181 ACCTTGGTGGTCCCTCCTGGCC-----TCAGTTTGAAGCTGGCTCCACAAGGACA 234  
QY 242 CGGTGACATGAGGGGAGCCACTGATCCATCTTCTGGCACCTTCTTCTCTGCTGCTTCT 301  
Db 235 CGGTGACATGAGGGGAGCCACTGATCCATCTTCTGGCACTTCTTCTCTGCTGCTTCT 294  
QY 302 CTCATATGGTGTGCCAGCTGTSCGGACACCCCTGTACCTGTCTTGGACACACCCCA 361  
Db 295 CTCATATGGTGTATTCAGCTGTGCCAGACCCCTGTGCTTGGACACACCCCA 354  
QY 362 GTCCCAACAGGGGTACCCCTGTGTGTGATGCTGTGGCTGTGTAAAGTGTGTGCAGC 421  
Db 355 GTCCCAACAGGGGTACCCCTGTGTGTGATGCTGTGGCTGTGTGCGAGTGTGTGCCG 414  
QY 422 GAGCTGGGGAGTCTCTGGACACCTGCATGTCTCGACCCCGACCCAGCCAGGCTGTGTG 481  
Db 415 GAGCTGGGGAGTCTCTGGACACCTGCATGTCTCGACCCCGACCCAGGCTGTGTGTTG 474  
QY 482 TCAGCTTGGGGAGCCCTGGCGCCATGGGGCTGTGTGTCTTCTTGGATGAGGATGACGG 541  
Db 475 TCAGCTTGGGGAGCCCGCCAGTGGCCGTGTGTGTGTCTTCTTGAAGAGGATGACGG 534  
QY 542 TAGCTGTAGGTGAATGGCGGAGGACCTGATGATGAGAGACCTTTAAACCCCAATTGCAG 601  
Db 535 GAGCTGTAGGTGAACGGCGCAGTACCTGTGATGGGAGACCTTTAAACCCCAATTGCAG 594  
QY 602 GTCTGTGCGCTGTGTGATGACGGTGGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTG 661  
Db 595 GGTTTGTGCGCTGTGTGATGACGGTGGTTCACCTGCTGCTGCTGCTGCTGCTGCTGCTG 654  
QY 662 CGGCTGCCAGCTGGGACTGCCACGCCCAAGAGAATACAGGTGCCAGGAAAGTGGCTG 721  
Db 655 CGGCTGCCAGCTGGGACTGCCACGCCCAAGAGAATACAGGTGCCAGGAGGTGGCTG 714  
QY 722 CCCCAGTGGGTATGTACAGGAGTGA---CACCAGGATCCAGGCTTCCAGCGGCA 778  
Db 715 CCCCAGTGGGTGTGTGACAGGAGTGTGTGACCGGCAATCCAGGCTTCCAGCGGCA 774  
QY 779 AGGACACCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 838  
Db 775 AGGACACCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 834  
QY 839 GAGCACAGCTGGGGCCCTGCTCAACCACTGTGGGCTGGGCAT 883  
Db 835 GAGCACAGCTGGGGCCCTGCTCAACCACTGTGGGCTGGGCAT 879

RESULT 6  
AX076919  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source  
I. .1266  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 216 a 418 c 390 g 242 t  
ORIGIN  
Query Match 49.2%; Score 434.6; DB 6; Length 1266;  
Best Local Similarity 79.9%; Pred. No. 1e-94;  
Matches 512; Conservative 0; Mismatches 129; Indels 0; Gaps 0;  
QY 243 GGTGACATGAGGGGAGCCCACTGATCCATCTTCTGGCCACTTCTCTGCTGCTGCTC 302  
Db 4 GGGGACATGAGAGGACACCCGAAGACCCCTCTCTGCTGCTGCTCTCTGCTGCTCCTC 63  
QY 303 TCAATGCTGTGTGCCAGCTGTGCCGACACACCTGTACCTGTCTTGGACACACCCCA 362  
Db 64 TCAAGGTGCTGTACCCAGCTGTGCCGACACACCTGTACCTGTGCTGCTGCTGCTC 123  
QY 363 TGCCCAACAGGGGTACCCCTGTGTGTGATGCTGTGGCTGTGTAAAGTGTGTGCAGG 422  
Db 124 TGCCCTGAGGATGACCCCTGTGTGTGATGCTGTGGCTGTGTGCTGCTGCTGCTGCTG 183  
QY 423 AGGCTGGGGAGTCTCTGCCACACCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 482  
Db 184 CGGCTGGGGAGCCCTTGCACCACTTCCACGCTGTGCGACGCGCCAGCCAGGCTGTGCT 243  
QY 483 CAGCTGGGGAGCCCTTGCAGCCCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 542  
Db 244 CAGCCCGGGAGGACCCGCTGTGCGGGGGCCCTGTGCTTGTGCGAGGACGACAGC 303  
QY 543 AGCTGTGAGGTGAATGGCGCGCAGGTACCTGGATGGAGAGACCTTTAAACCCCAATTGCAG 602  
Db 304 AGCTGTGAGGTGAACGGCGCTGTATCGGAAGGGGAGACCTTCCAGCCCACTGCTGAG 363  
QY 603 GTCTGTGCGCTGTGATGACGCTGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 662  
Db 364 ATCCGCTGCGCTGTGAGGAGCGGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 423  
QY 663 CGGCTGCCAGCTGGGACTGCCACGCCCAAGAGAATACAGGTGCCAGGAAAGTGTGCTG 722  
Db 424 CGGCTGCCAGCTGGGACTGCCACGCCCAAGAGAATACAGGTGCCAGGAGGTGCTGCTG 483  
QY 723 CCGAGTGGGTATGTGACAGGAGTGCACCGGGGATCCAGCGCTTCCAGCGGCGCAAGA 782  
Db 484 CCGTGTGGGTGTGCGGCCAAGGAGGGGACTGGGGACCCAGCCCTTCCAGCCCAAGGA 543  
QY 783 CACCAACTTCTGCGCTTGTCTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 842  
Db 544 CCGCAGTTTCTGCGCTTGTCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 603  
QY 843 ACAGCTGGGGCCCTGCTCAACCACTGTGGCTGGGCAT 883  
Db 604 AGGCTGGGGACCCCTGCTCGACCACTGTGGCTGGGCAT 644

Dd

604 ACGGCTGGGGACCCCTGCTGCACCACCTGTGGCTGGGCAT 644

RESULT 8  
AF083500 linear PRI 04-NOV-1998  
LOCUS AF083500 1283 bp mRNA growth factor-like protein  
DEFINITION Homo sapiens connective tissue growth factor-like protein precursor, mRNA, complete cds.  
ACCESSION AF083500  
VERSION AF083500.1 GI:3462835  
KEYWORDS .  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1 (bases 1 to 1283)  
AUTHORS Kumar,S., Hand,A.T., Connor,J.C., Dodds,R.A., Ryan,P.J., Trill,J.G., Fisher,S.M., Slemmon,J.R., Lipshutz,D.B., Bartholomew,V., James,I.E., Riemann,D.J., Gowen,M. and Lee,J.C.  
TITLE Identification and cloning of CTGF-L from human osteoblasts, a novel cysteine rich protein containing an IGF binding domain  
JOURNAL Bone 23 (5), 5240 (1998)  
REFERENCES 2 (bases 1 to 1283)  
AUTHORS Kumar,S.  
TITLE Direct Submission  
JOURNAL Submitted (11-AUG-1998) Bone & Cartilage Biology, UW 2109,  
SmithKline Beecham, 709 Swedeland Road, King of Prussia, PA 19406,  
USA  
FEATURES  
source Location/Qualifiers  
1..1283  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="20"  
/map="20q12-q13"  
9..761  
/cell\_type="primary osteoblast"  
/note="CTGF-L; encodes IGF binding (IGFB), Von Willebrand factor type C (WVC) and thrombospondin type I (TSP1) domains; member of the CCN (CTGF/cyr61/nov) family; lacks the fourth carboxy-terminal domain present in other members of the CCN family"  
/codon\_start=1  
/product="connective tissue growth factor-like protein precursor"  
/protein\_id="AAC70350.1"  
/db\_xref="GI:3462836"  
/translation="MRGTPKTHLLAFSLILCLLSKVRTOLCPTPCTCPWPppRCPPLGV LVIDSGRCRVCARLGEPCDQHLVCDASQGLVCPAGPGGRGALCLLAEDDSCEV NGRLREGTFPHFQIRCRDEGGFTCVPLCSDEDVLPFSWCDCPIPRRVEVLGKKCEP WYCGQGGLGTQFLPAQGFSGLVSSLPVPVGPPEWSLTAWGSPCTTCGLGWATRVS QNRFCLETQRLLCLRSPCPGRSPNSAP"  
BASE COUNT 235 a 418 c 389 g 241 t  
ORIGIN  
  
Query Match 49.2%; Score 434.6; DB 9; Length 1283;  
Best Local Similarity 79.9%; Pred. No. 1e-94;  
Matches 512; Conservative 0; Mismatches 129; Indels 0; Gaps 0;  
  
QY 243 GGTGACATGAGGGAGGCCACTGATCCATCTCTGGCCACTTCTCTGCTGCTGCTTC 302  
Db 3 GGGACATGAGGACACCGAAGACCCCTCTGGCCCTTTCTCCCTCTCTGCTGCTCTC 62  
  
QY 303 TCAATTGGTGTCGCCAGCTGTGCCGGACACCCCTGTACCTGTGCTTGGAGACACCCCAG 362  
Db 63 TCAAAGGTGCGTAGCCAGCTGTGCCGACACCATGTACCTGCCCTCTGGCACCTCCCGCA 122  
  
QY 363 TGCCACACAGGGGTACCCCTGCTGCTGGATGCTGTGGCTGCTGTAAACTGTGTGACCG 422  
Db 123 TGCCCGCTGGGASTACCCCTGTGCTGGATGCTGTGGCTGCTGTGGCGGTGTGTGACCG 182  
  
QY 423 AGGCTGGGGAGTCTCGCACACCTGCATGTCTGCAGCCCCAGCCAGGCGCTGTTGT 482  
Db 183 CGGCTGGGGAGCCTTCGACCACTCCACGCTCTCGACGCCAGCCAGCCAGGCTGCTGCTGC 242

|    | Query Match           | 49.2%   | Score 434.6;     | DB 6;     | Length 1293; |
|----|-----------------------|---|------------------|-----------|--------------|
|    | Best Local Similarity | 79.9%;  | Pred. No. 1e-94; |           |              |
|    | Matches 512;          | Conservative 0;   | Mismatches 129;  | Indels 0; | Gaps 0;      |
| QY | 243                   | GGTGACATGAGGGCAGCCACATGATCCATCTTCTGGCCACTTCCTCTCCTCTGCCCTCTC  | 302              |           |              |
| Db | 16                    | GGGACATGAGGGCACACCGAAGACCACCTCTGGCCTTCCTCTCCTCTGCCCTCTC       | 75               |           |              |
| QY | 303                   | TCAATGTGTGTGCCAGCTGTGCCGGACACCCCTGTACCTTGCTCTTGGACACACACCCACG | 362              |           |              |
| Db | 76                    | TCAAAGGTGCGTACCAGCTGTGCCGGACACCATGTAGCTGCCCTGGCCACCTCCCGA     | 135              |           |              |
| QY | 363                   | TGCCCCACAGGGGTACCCCTGTGTGGATGGCTGTGGCTGTGAAGTGTGTGCACGG       | 422              |           |              |
| Db | 136                   | TGCCCGCTGGGAGTACCCCTGTGTGGATGGCTGTGGCTGTGCCGGGTATGTGCACGG     | 195              |           |              |
| QY | 423                   | AGGCTGGGGAGTCCTGCGACCACTCTGCATGTCTGCGACCCACGAGGCGCTGGTTGT     | 482              |           |              |
| Db | 196                   | CGGCTGGGGAGCCCTGCGNCAACTCCAGCTGTGCAGCCGACACGAGGCGCTGTCTGC     | 255              |           |              |
| QY | 483                   | CAGCCTTGGGCGAGGGCCCTGGCGGCCATGGGGCTGTGTGTCTTTGGATGAGGATGACGGT | 542              |           |              |

|    | Query Match           | 49.28;   | Score 434.6;     | DB 6;     | Length 1293; |
|----|-----------------------|--|------------------|-----------|--------------|
|    | Best Local Similarity | 79.9%;   | Pred. No. 1e-94; |           |              |
|    | Matches 512;          | Conservative 0;  | Mismatches 129;  | Indels 0; | Gaps 0;      |
| Qy | 243                   | GGTGACATGAGGGCAGCCCACTGATCCATCTTCTGGCCACTTCCTTCCTCTGCCTTCCTC | 302              |           |              |
| Db | 1278                  | GGGGACATGAGGACACACCGAAGACCCACTCTGGCCCTTCTCCCTTCCTCTGCCCTCCTC | 1219             |           |              |
| Qy | 303                   | TCAATGGTGTGTGCCAGCTGTGCGGACACCCCTGTACCTGTCTTGGACACACACCCCG   | 362              |           |              |
| Db | 1218                  | TCAAGTGGGTACCCAGCTGTGCGCGACACCATGTACCTGCCCTGGCCACCTCCCGA     | 1159             |           |              |
| Qy | 363                   | TGCCACAGGGGTACCCCTGGTCTGGATGGCTGTGGCTGTAAAGTGTGTGCACGG       | 422              |           |              |
| Db | 1158                  | TGCCCGCTGGAGTACCCCTGGTCTGGATGGCTGTGGCTGTGCCGGGTATGTGCACGG    | 1099             |           |              |
| Qy | 423                   | AGGCTGGGGAGTCTCGGACACCTGCATGTCTGGACCCCGACCGCCCTGGTTGT        | 482              |           |              |
| Db | 1098                  | CGGCTGGGGAGCCCTCGGACCACTCCAGCTCTGGACGCCACGCCAGGGCCCTGGTCTGC  | 1039             |           |              |
| Qy | 483                   | CAGCTGGGCGAGCCCTCGCGCCATGGGGCTGTGTCTCTTGGATGAGGATGACGGT      | 542              |           |              |
| Db | 1038                  | CAGCCCCGGGCGAGACCCGGTGGCCGGGGGGCCCTGTGCCCTCTTGGCAGAGGACGACGC | 979              |           |              |
| Qy | 543                   | AGCTGTGAGGTGAATGGCCCGAGTGACTCGATGGAGAGACCTTTAAACCCAATTTGCAGG | 602              |           |              |

```

|||||
978 AGCTGTAGGTGAACGGCGCGCTGTATCGGAAGGGAGACCTTCACGCCCTACTGTCAGC 919
603 GTCCCTGTGCGGTGTGATGACGGTGGCTTCACTCCCTCGCGCTGTGAGTGAGGATGTG 662
918 ATCCCTCGCGGTGTGAGAGCGCGGTTCACCTCGCTGCGCTGTGAGCGAGGATGTG 859
663 CGGCTGCCAGTGGGACTGCCACGCCGCCAAGAGAAATACAGGTGCCAGGAAGTGTCTGC 722
858 CGGCTGCCAGTGGGACTGCCACGCCGCCAAGAGAGGTGCGAGGTCTGGCAAGTGTCTGC 799
723 CCCGAGTGGGTATGTACAGGAGGTGACACGGCGGATCCAGCGGTCTCCACGCCGAAGA 782
798 CCGTGTGGTGTGCGGCAAGAGGGGACTGGGGACCCAGCCCTTCCAGCCCAAGGA 739
783 CACCAACTTTCCTGCTGTACTCTGCTCTGCTGTGATGCTCTGTGCTTGTCCAAATTTGGAC 842
738 CCCAGTGTTCGCGCTGTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 679
843 ACAGCTGGGGCCCTGCTCAACACCTGTGGGCTGGGCAT 883
678 ACGGCTGGGACCTGCTCGACACCTGTGGGCTGGGCAT 638

```

## RESULT 11

```

AF074604
LOCUS
DEFINITION
Homo sapiens connective tissue growth factor-related protein
precursor (CT58) mRNA, complete cds.
AF074604
VERSION
KEYWORDS
SOURCE
Homo sapiens.
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1309)
Rowles,J. and Gendler,S.
CT58, a new member of the connective tissue growth factor family,
interacts with the breast cancer associated mucin MUC1
Unpublished
2 (bases 1 to 1309)
Rowles,J. and Gendler,S.
Direct Submission
Submitted (25-JUN-1998) Biochemistry and Molecular Biology, Mayo
Clinic Scottsdale, 13400 E. Shea Blvd., Scottsdale, AZ 85259, USA
Location/Qualifiers
1..1309
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="HeLa"
1..1309
/gene="CT58"
7..759
/gene="CT58"
/Note="contains three of four modules found in growth
regulators related to connective tissue growth factor as
described in FEBS Letters 327:125-130,1993; identified in
yeast two-hybrid screen using the epithelial mucin MUC1 as
bait"
/codon_start=1
/product="connective tissue growth factor-related protein
precursor"
/db_xref="GI:3328192"
/translation="MRGTPKTHLAFSLLLSKVRLCTPTCTCPWPPRPLGLVP
LYDCCGCRVCRARLGPDLHVCDSQGLVCPGAGPGGRGALCLLADSSCEV
NGRLYREGTQPHCSIRCEGDGFTCPVLCSDVRPLSDWCPHPRVRLGKCCPE
WVGQGLGTQPLPAQGPFGSLVSLPPGVPCEWSTANGPCSTTCLGLMATRVSN
QNRFLRQLRCLSRCPSPSRGSPNSAF"

```

## gene

## CDS

```

BASE COUNT
ORIGIN
261 a 418 c 387 g 242 t 1 others

```

```

Query Match 49.2%; Score 434.6; DB 9; Length 1309;
Best Local Similarity 79.9%; Pred. No. 1e-94;
Matches 512; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

QY 243 GGTGACATGAGGGGAGAGCCCACTGATCCATCTTGGGCCACTTCTCTCTGCTTCTC 302
DB 1 GGGGACATGAGAGGAGGACACACCAAGACCCACCTCTTGGGCTTCTCTCTCTCTC 60
QY 303 TCAATGTGTGTGCGCCAGCTGTGCGGGACACCTGTACTGTCTTGTGACACACCCAG 362
DB 61 TCAAGGTGTGTGCGCCAGCTGTGCGGGACACCATGTACTGTCTCTGCGCCCTGCCGA 120
QY 363 TCCACACAGGGGTGTACCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 422
DB 121 TCCCGCTGGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
QY 423 AGGCTGGGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 482
DB 181 CGGCTGGGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 240
QY 483 CAGCTGGGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 542
DB 241 CAGCTGGGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 300
QY 543 AGCTGTGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 602
DB 301 AGCTGTGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 360
QY 603 GTCCGTGTCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 562
DB 361 ATCCGCTGCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 420
QY 663 CGGCTGCCAGTGGGACTGCCAGCCCAAGAGAAATACAGGTGCCAGGAAGTGTCTGC 722
DB 421 CGGCTGCCAGTGGGACTGCCAGCCCAAGAGAGGTGTGTGTGTGTGTGTGTGTGT 480
QY 723 CCCGAGTGGGTATGTGACAGGAGTGTGACACGGCGATCCAGCGCTTCCAGCGCGCAAGA 782
DB 481 CTTGAGTGGGTGTGCGGCAAGGAGGGGACTTGGGGACCCAGCCCTTCCAGCCCAAGA 540
QY 783 CACCAACTTTCCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 842
DB 541 CCCAGTGTTCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
843 ACAGCTGGGGCCCTGCTCAACACCTGTGGGCTGGGCAT 883
601 ACGGCTGGGAGCTGTCTCGACACCTGTGGGCTGGGCAT 641

```

## RESULT 12

## AF100780

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

```

AF100780
Homo sapiens connective tissue growth factor related protein WISP-2
(WISP2) mRNA, complete cds.
AF100780.1 GI:4028582
Homo sapiens.
Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Pennica,D., Swanson,T.A., Welsh,J.W., Roy,M.A., Lawrence,D.A.,
Lee,J., Brush,J., Taneyhill,L.A., Deuel,B., Lew,M., Watanabe,C.,
Cohen,R.L., Mellem,M.F., Finley,G.G., Quirke,P., Goddard,A.D.,
Hillan,K.J., Guiney,A.L., Botstein,D. and Levine,A.J.
WISP genes are members of the connective tissue growth factor
family that are up-regulated in wnt-1-transfected cells and
aberrantly expressed in human colon tumors
Proc. Natl. Acad. Sci. U.S.A. 95 (25), 14717-14722 (1998)
99061933
9843955
2 (bases 1 to 1427)

```

**AUTHORS** Pennica, D.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (23-OCT-1998) Molecular Oncology, Genentech, Inc., 1 DNA  
Way, South San Francisco, CA 94080, USA  
**FEATURES** Location/Qualifiers

```

1. .1427
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="20"
/map="20q12-q13.1"
/tissue_type="lung"
/dev_stage="fetus"
1. .1427
/gene="WISP2"
136. .908
/gene="WISP2"
/codon_start=1
/product="connective tissue growth factor related protein
WISP-2"
/protein_id="AAC96322.1"
/db_xref="GI:4028583"
/translation="MRGTPKTHLAFSLILLSKVETQLCPCTCPWPPRCPLGVP
LVLDGCCRCVRCARIGECFDQLHVCDSQGLVCQAGFGGAGTCLLAEDSSCEV
NKLRYHEGTFFPHCSIRKCEDEGGTVPFCSDEVDRLPSWDCPHPRRVVEVLKKCCPE
WYCGCGLEGTQPLPAQGPQFGLVSLPLPPGVPCPEWSTAWGFCSTCGIGMATRVSN
QNRFGGRTQRLCSRPCPPRRGRSPQNSAF"
261 a 468 c 417 g 280 t 1 others
BASE COUNT
ORIGIN

```

|                       |              |              |                 |              |
|-----------------------|--------------|--------------|-----------------|--------------|
| Query Match           | 49.2%        | Score 434.6; | DB 9;           | Length 1427; |
| Best Local Similarity | 79.9%;       |              |                 |              |
| Matches 512;          | Conservative | 0;           | Mismatches 129; | Indels 0;    |
|                       |              |              |                 | Gaps 0;      |

|     |    |  |     |
|-----|----|--|-----|
| 243 | QY | GGTGACATGAGGGGAGCCCACTGATCCATCTTCTGGCCACATTCCTTCTCTCTGCTTCTC   | 302 |
| 150 | Db | GGGGACATGAGAGGACACCCGAAGAGCCCACTCTCTGGCCCTTCTCCCTCTCTGCTCTCT   | 209 |
| 303 | QY | TCAATGGTGTGCCCAGCTGSCGGAGACACCTGTACCTGTCTCTTGGACACACCCACG      | 362 |
| 210 | Db | TCAAGGTGCTACCAAGCTGTGCCCGACACCAATGTACCTGCCCTTGGCCACCTCCCGA     | 269 |
| 363 | QY | TGCCCCAGGGGTACCCCTGTGCTGGATGGCTGTGGCTGTAAAGTGTGTGCACGG         | 422 |
| 270 | Db | TGCCCGCTGGAGTACCCCTGTGTCTGGATGGCTGTGGCTGTGCCGGTATGTGCACGG      | 329 |
| 423 | QY | AGGCTGGGGAGTCTTCGCACCACTTCATGTCTGGACCCCAAGCCAGGGCCCTGGTTGT     | 482 |
| 330 | Db | CGGCTGGGGAGCCCTGCGACCAACTCCAGTCTGGCAGCCAGCAGGGCCCTGGTCTGC      | 389 |
| 483 | QY | CAGCCTGGGACGGCCCTGCGGCCATGGGGCTGTGTCTCTTGGATGAGGATGACGGT       | 542 |
| 390 | Db | CAGCCCGGGCAGGACCCGGTGGCCGGGGGGCCCTGTGCCCTCTTGGCAGAGGACGACGC    | 449 |
| 543 | QY | AGCTGTGAGGTGAATGGCCGACAGTACTCGATGGAGAGACCTTTAAACCCAAATTGCAGG   | 602 |
| 450 | Db | AGCTGTGAGGTGAACGGCCGCTGTATCGGGAAGGGAGACCTTCCAGCCCCACTGCAGC     | 509 |
| 603 | QY | GTCTGTGGCTGTGATGACGGTGGCTTCACTCTGCCCTGGCGCTGTGCATGAGGATGTG     | 662 |
| 510 | Db | ATCCGCTGCCGCTGCGAGGACGCGGCTTCACTGTGGCTGCCGTGTGCACGCGAGGATGTG   | 569 |
| 663 | QY | CGGCTGCCAGCTGGGACTGCCACGCCCCAAGAGAATACAGTGTCCAGGAAAGTGTCTGC    | 722 |
| 570 | Db | CGGCTGCCAGCTGGGACTGCCCCCAAGGAGGCTCGAGGTCTCTGGGCAAGTGTCTGC      | 629 |
| 723 | QY | CCCGAGTGGGTATGTGACCAAGGAGTGCACACCGCGCATCCAGCGCTCCAGCGGCGCAAGGA | 782 |
| 630 | Db | CCTGAGTGGGTGTGCGGCCAAGGAGGGGACTGGGGACCCAGCCCTTCCAGCCCCAAGGA    | 689 |
| 783 | QY | CACCAACTTCTGCCCTTGTCACTCTCTGCTCTGTCTGTATGCTCTCTGTCCAAATTGGAGC  | 842 |
| 690 | Db | CCCGACTTTCTGGCCCTGTCTCTCTCTGCCCTTGGTGTGCCCTGCCAGATGAGC         | 749 |

[illegible]

ACCESSION BC017782  
VERSION BC017782.1 GI:17389482  
KEYWORDS MGC.

| ORGANISM  | Homo sapiens  |
|-----------|---|
| REFERENCE | 1 (bases 1 to 1450)   |
|           | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.        |
|           | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; |

**JOURNAL TITLE** Direct Submission  
Submitted (03-DEC-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

COMMENT

Contact: MGC help desk [help777@genome.ucsf.edu](mailto:help777@genome.ucsf.edu)  
Email: [cqabps@email.nih.gov](mailto:cqabps@email.nih.gov)  
Tissue Procurement: CLONTECH  
cDNA Library Preparation: CLONTECH Laboratories, Inc.  
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)  
Center, Stanford University School of Medicine, Stanford, CA 94305  
Web site: <http://www-shgc.stanford.edu>  
Contact: (Dickson, Mark) [mcd@paxil.stanford.edu](mailto:mcd@paxil.stanford.edu)  
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAL Plate: 36 Row: m Column: 3  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA qi: 4507922.

```

FEATURES
Location/Qualifiers
1..1450
/organism="Homo sapiens"
/db_xref="LocusID:8839"
/db_xref="taxon:9606"
/clone="MGC:22271 IMAGE:4691574"
/tissue_type="Lung"
/clone_lib="NIH_MGC_77"
/lab_host="DH10B"
/notes="Vector: pDNR-LIB"
165...917
CDS

```

```

CDS
165...917
/codon_start=-1
/product="wnt1 inducible signaling pathway protein 2"
/protein_id="AAH1782.1"
/db_xref="GI:17389483"
/translation="MRGTPKTHLAFSLCLLSKVRTQLCPTCTCPWPPRCPLGVP
LVLDGCCRVACRRRIGPCDQLHVCDDASGLVCOPCGGGRGALCLIAEDSSCE
NGRLYRGETFQPHCSIRCRCEGGGTCVPLCSEDVRLPSWDCPHRRRVYELKCKQE
WVCQGGGLGTQLPAGQFGVSLVSLPGVPCPEWSTAWGFCSTTCGUGMAIRVSN
NFRFRLTQRRLCSLPPSPGRSPQNSAF"
272 a 457 c 441 g 280 t
BASE COUNT
ORIGIN

```

|                       |              |                  |                 |              |
|-----------------------|--------------|------------------|-----------------|--------------|
| Query Match           | 49.2%        | Score 434.6;     | DB 9;           | Length 1450; |
| Best Local Similarity | 79.9%;       | Pred. No. le-94; |                 |              |
| Matches 512;          | Conservative | 0;               | Mismatches 139; | Indels 0;    |
|                       |              |                  |                 | Gaps 0;      |

| QY | 243 | GGTGACATGAGGGGACGCCACTGATCCATCTTCTGGCCACTTCCCTTCCCTGCGCTCTC | 302 |
|----|-----|---|-----|
| QY | 159 | GGGACATGAGAGGCACACCGAAGACCCACCTCTGGCGTTCTCCCTCTCTGCGCTCTC   | 218 |
| Db | 159 | GGGACATGAGAGGCACACCGAAGACCCACCTCTGGCGTTCTCCCTCTCTGCGCTCTC   | 218 |



```
Db 126 TGCCGCTGGGAGTACCCCTGGTGGTGATGGCTGTGGCTGCCGGGTATGTGCACGG 185
QY 423 AGGCTGGGGAGTCTCTCGACACCTGCTCGACCCAGCCAGGGCTGTGTGT 482
Db 186 CGGCTGGGGAGCCCTGGACCAACTCCACGCTCGCGACGCCAGCCAGGGCTGTGTGC 245
QY 483 CAGCCTGGGGAGGGCCCTGGCGGCATYGGGGCTGTGTCTCTTGGATGAGGATGACGGT 542
Db 246 CAGCCGGGGAGGACCCGGTGGCGGGGGCCCTGTGCCTCTTGGCAGAGGAGACACAGC 305
QY 543 AGCTGTGAGGTGAATGGCCGACAGGTACCTGATGGAGAGACCTTTAAACCCAAATGCAAG 602
Db 306 AGCTGTGAGGTGAACGGCCGCTGTATCGGAGGGGAGACCTTCCAGCCCACTGCACG 365
QY 603 GTCTGTGCCGCTGTGATGACGGTGGCTTTCACCTGCCCTGCGGCTGTGCAGTGCAGGATGTG 662
Db 366 ATCCGCTGCCGCTGGGAGGACGGCGCTTACCTGCGTGGCGCTGTGCAGCGAGGATGTG 425
QY 663 CGGCTGCCAGCTGGGACTGCCACGGCCCGCCCAAGAGATACAGGTGCCAGGAAAGTGTGC 722
Db 426 CGGCTGCCAGCTGGGACTGCCCGCCACCCAGGAGGTTCGAGGTCTTGGGCAAGTGTGTGC 485
QY 723 CCCGAGTGGGTATGTGACCCAGGAGTGACACCGGGCATCCAGCGCTCCACGGCGCAAGGA 782
Db 486 CCTGAGTGGGTGTGGGCCCAAGAGGGGGACTGGGGACCGCCCTTCCA--GCCCAAGGA 543
QY 783 CACCAACTTTCTGCCCTTGTCTCAGTCTGCTGATGCTCTTGTCCAAATTTGGAGC 842
Db 544 CCCCAGTTTCTGGCCCTGTCTCTTCCCTGCCCTGGTGTCCCTGCCAGAAATGGAGC 603
QY 843 ACAGCCTGGGGCCCTGTCTCAACCACTGTGGGTGGGCAT 883
Db 604 ACGGCTGGGGACCTGTCTCGACCACTGTGGGCTGGGCAT 644
```

Search completed: July 29, 2003, 00:48:29  
Job time : 2456.08 secs





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 23, 2003, 09:03:08 ; Search time 233.823 Seconds  
(without alignments)  
8504.365 Million cell updates/sec

Title: US-10-010-408-1-copy\_1\_883

Perfect score: 883

Sequence: 1 GAGGCTTCGATCTCCAGAG.....ACCACCTGCGGTGGGCAT 883

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 segs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_101002.\*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*
- 24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Query Match | Score | Length | ID | Description |
|------------|-------------|-------|--------|----|-------------|
| 1          | 883         | 100.0 | 1708   | 20 | AAZ07516    |
| 2          | 727.6       | 82.4  | 1734   | 20 | AAZ07516    |
| 3          | 635         | 71.9  | 753    | 20 | AAZ07517    |
| 4          | 566         | 64.1  | 681    | 20 | AAZ07521    |
| 5          | 553.2       | 62.7  | 753    | 20 | AAZ07521    |
| 6          | 460.6       | 52.2  | 1522   | 20 | AAZ07521    |
| 7          | 440         | 49.8  | 1337   | 22 | AAZ07521    |
| 8          | 439.8       | 49.8  | 1352   | 22 | AAZ07521    |
| 9          | 434.6       | 49.2  | 1266   | 21 | AAZ07521    |

|    |       |      |       |    |          |                    |
|----|-------|------|-------|----|----------|--------------------|
| 10 | 434.6 | 49.2 | 1266  | 22 | AAZ07521 | Human cDNA sequenc |
| 11 | 434.6 | 49.2 | 1266  | 22 | AAZ07521 | PRO261 coding sequ |
| 12 | 434.6 | 49.2 | 1266  | 22 | AAZ07521 | Human angiogenesis |
| 13 | 434.6 | 49.2 | 1267  | 21 | AAZ07521 | Human PRO261 cDNA  |
| 14 | 434.6 | 49.2 | 1285  | 19 | AAZ07521 | Human connective t |
| 15 | 434.6 | 49.2 | 1293  | 20 | AAZ07521 | Human WISP-2 prote |
| 16 | 434.6 | 49.2 | 1309  | 22 | AAZ07521 | Connective tissue  |
| 17 | 430.2 | 48.7 | 750   | 20 | AAZ07521 | Human WISP-2 prote |
| 18 | 430.2 | 48.7 | 1257  | 20 | AAZ07521 | EGF-like homologue |
| 19 | 423   | 47.9 | 738   | 20 | AAZ07521 | Human WISP-2 prote |
| 20 | 420.6 | 47.6 | 841   | 20 | AAZ07521 | Human WISP-2 prote |
| 21 | 210   | 23.8 | 210   | 20 | AAZ07519 | Rat HICP IGFBP dom |
| 22 | 203.4 | 23.0 | 2136  | 22 | AAZ07518 | Human full-length  |
| 23 | 177   | 20.0 | 177   | 20 | AAZ07518 | Rat HICP VWC domai |
| 24 | 169.8 | 19.2 | 586   | 22 | AAZ07518 | Human foetal liver |
| 25 | 169.8 | 19.2 | 586   | 22 | AAZ07518 | Probe #6698 for ge |
| 26 | 169.8 | 19.2 | 586   | 22 | AAZ07518 | Human brain expres |
| 27 | 169.8 | 19.2 | 586   | 22 | AAZ07518 | Human bone marrow  |
| 28 | 169.8 | 19.2 | 586   | 22 | AAZ07518 | Probe #6485 for ge |
| 29 | 169.8 | 19.2 | 586   | 22 | AAZ07518 | Probe #8254 used t |
| 30 | 169.8 | 19.2 | 586   | 22 | AAZ07518 | Human genome-deriv |
| 31 | 169.8 | 19.2 | 13255 | 22 | AAZ07518 | Human immune/haema |
| 32 | 139.8 | 15.8 | 199   | 22 | AAZ07518 | Human foetal liver |
| 33 | 139.8 | 15.8 | 199   | 22 | AAZ07518 | Probe #16578 for g |
| 34 | 139.8 | 15.8 | 199   | 22 | AAZ07518 | Human brain expres |
| 35 | 139.8 | 15.8 | 199   | 22 | AAZ07518 | Human bone marrow  |
| 36 | 139.8 | 15.8 | 199   | 22 | AAZ07518 | Probe #15682 for g |
| 37 | 139.8 | 15.8 | 199   | 22 | AAZ07518 | Probe #21336 used  |
| 38 | 139.8 | 15.8 | 199   | 22 | AAZ07518 | Human genome-deriv |
| 39 | 128.4 | 14.5 | 2075  | 16 | AAZ07518 | Connective tissue  |
| 40 | 128.4 | 14.5 | 2075  | 18 | AAZ07518 | Human connective t |
| 41 | 128.4 | 14.5 | 2075  | 18 | AAZ07518 | Connective tissue  |
| 42 | 128.4 | 14.5 | 2075  | 19 | AAZ07518 | Human connective t |
| 43 | 128.4 | 14.5 | 2075  | 20 | AAZ07518 | Human connective t |
| 44 | 128.4 | 14.5 | 2075  | 21 | AAZ07518 | Human connective t |
| 45 | 128.4 | 14.5 | 2075  | 21 | AAZ07518 | Human connective t |

ALIGNMENTS

RESULT 1

AAZ07516  
ID AAZ07516 standard; cDNA; 1708 BP.

XX AAZ07516;

XX 26-NOV-1999 (first entry)

DT Rat HICP polypeptide encoding cDNA.

XX Heparin-induced CCN-like protein; HICP; cell-associated activity; ss;  
cardiovascular disorder; aberrant cell proliferation; fibrotic disorder.

OS Rattus sp.

PN WO9947556-A2.

XX 23-SEP-1999.

XX 18-MAR-1999; 99WO-US05999.

XX 19-MAR-1998; 98US-0044273.

XX (TUFT ) TUFTS COLLEGE.

XX Castellot JJ;

XX WPI; 1999-562060/47.

XX P-PSDB; AAY27434.

XX Nucleic acid sequences encoding rat heparin-induced CCN-like protein,  
used in methods to identify modulators or in diagnostic applications

|    |   |  |
|----|---|--|
| XX | Claim 2: Fig 1: 108pp; English.   |  |
| PS | This cDNA encodes a rat heparin-induced CN-like protein (HICP) protein.   |  |
| XX | Agents that stimulate or inhibit HICP protein activity or expression.     |  |
| CC | antisense HICP nucleic acid molecules and HICP antibodies, can be used to |  |
| CC | modulate cell-associated activity. HICP modulators can be used to treat   |  |
| CC | disorders characterized by aberrant HICP protein activity or expression.  |  |
| CC | Probes capable of hybridizing to HICP mRNA or antibodies specific for     |  |
| CC | HICP can be used to detect HICP activity in a biological sample. HICP     |  |
| CC | can be used to treat disorders, such as a cardiovascular or fibrotic      |  |
| CC | disorder, characterized by aberrant cell proliferation.                   |  |
| XX |   |  |
| SQ | Sequence 1708 BP; 362 A; 486 C; 478 G; 382 T; 0 other;                    |  |
|    | Query Match 100.0%; Score 883; DB 20; Length 1708;                        |  |
|    | Best Local Similarity 100.0%; Pred No. 6, 7e-235;                         |  |
|    | Matches 883; Conservative 0; Mismatches 0; Indels 0; Gaps 0;              |  |
| QY | 1 GACGCTTCGATCTCCAGAGGACCTGGGTGGGACAGGGGCTTGGCAAGCTCGAGCC 60              |  |
| Db | 1 GACGCTTCGATCTCCAGAGGACCTGGGTGGGACAGGGGCTTGGCAAGCTCGAGCC 60              |  |
| QY | 61 GCTGGCAGTGGCTTGGATGGAGTCTTTATCTAGTGGAACTGAGGAGCTTAAGAGGCTC 120         |  |
| Db | 61 GCTGGCAGTGGCTTGGATGGAGTCTTTATCTAGTGGAACTGAGGAGCTTAAGAGGCTC 120         |  |
| QY | 121 CTGTACAGTGTCTTAAAGCTTAGACACTGTGTGGCTTGGCTTCCACACACTGTTCAGA 180        |  |
| Db | 121 CTGTACAGTGTCTTAAAGCTTAGACACTGTGTGGCTTGGCTTCCACACACTGTTCAGA 180        |  |
| QY | 181 CACCTTCGTGGTGGGCTCCACGGCCCTACCTTCAGGTTTGAAGCTGGCTTCCACAGGGAC 240      |  |
| Db | 181 CACCTTCGTGGTGGGCTCCACGGCCCTACCTTCAGGTTTGAAGCTGGCTTCCACAGGGAC 240      |  |
| QY | 241 ACGGTGACATGAGGGCAGCCACCTAGTATCCATCTCTGGCCACTTCTTCCTCTGCGCTTC 300      |  |
| Db | 241 ACGGTGACATGAGGGCAGCCACCTAGTATCCATCTCTGGCCACTTCTTCCTCTGCGCTTC 300      |  |
| QY | 301 TCTCAATGTTGTGCCAGCTGTGCCGACACCTGTACCTGTCTTGGACACACACCC 360            |  |
| Db | 301 TCTCAATGTTGTGCCAGCTGTGCCGACACCTGTACCTGTCTTGGACACACACCC 360            |  |
| QY | 361 AGTGCCACAGGGGGTACCCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 420       |  |
| Db | 361 AGTGCCACAGGGGGTACCCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 420       |  |
| QY | 421 GGAGGCTGGGGAGTCTCGACACCTGCATGTCTGCCACCCACGAGGCGCTGGTTT 480            |  |
| Db | 421 GGAGGCTGGGGAGTCTCGACACCTGCATGTCTGCCACCCACGAGGCGCTGGTTT 480            |  |
| QY | 481 GTACGCTGGGGCAGGCTTGGCGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 540      |  |
| Db | 481 GTACGCTGGGGCAGGCTTGGCGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 540      |  |
| QY | 541 GTAGCTGTGAGTGAATGCGCGAGTACCTGGATGGAGAGACCTTTAAACCAATTGCA 600          |  |
| Db | 541 GTAGCTGTGAGTGAATGCGCGAGTACCTGGATGGAGAGACCTTTAAACCAATTGCA 600          |  |
| QY | 601 GGGTCTGTGCGCTGTGATGACGGTGGCTTACCTGGCTTGGCTGGCTGGCTGGCTGGCTGG 660      |  |
| Db | 601 GGGTCTGTGCGCTGTGATGACGGTGGCTTACCTGGCTTGGCTGGCTGGCTGGCTGGCTGG 660      |  |
| QY | 661 TCGCGCTGCCAGCTGGGACTGCCACGCCCAAGAGATACAGTGCAGGAAAGTGC 720             |  |
| Db | 661 TCGCGCTGCCAGCTGGGACTGCCACGCCCAAGAGATACAGTGCAGGAAAGTGC 720             |  |
| QY | 721 GCGCCGAGTGGGTATGTACACAGGAGTGCACACCGGCGATCCAGCGCTCCACGGCGCAAG 780      |  |
| Db | 721 GCGCCGAGTGGGTATGTACACAGGAGTGCACACCGGCGATCCAGCGCTCCACGGCGCAAG 780      |  |
| QY | 781 GACACCAACTTTCTGCGCTTGTGACTCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840   |  |
| XX |   |  |

### Query Match

82.48; Score 727.6; DB 20; Length 1734;

## Len

3 20;

27.6

50

8

cy Ma

— 450 —

1111

0776

**II**

1111

1111

1111  
TOWN

5-40



|    |     |  |     |
|----|-----|--|-----|
| Db | 421 | CCGAGCTGGGACTGCCCAACGCCCAAGAGAAATACGGTGCCAGGAAGTGTCTGCCCCGAG | 480 |
| Qy | 729 | TGGGTATGTGACACAGGGAGTGACACCGCGCATCCAGCGTCCACGCGCAAGGACACCAA  | 788 |
| Db | 481 | TGGGTATGTGACACAGGGAGTGACACCGCGCATCCAGCGTCCACGCGCAAGGACACCAA  | 540 |
| Qy | 789 | CTTTCTGCCCTTGTCACCTCCTCTGCTGATGCTCTTGTCCTCAATTTGGAGCACAGCC   | 848 |
| Db | 541 | CTTTCTGCCCTTGTCACCTCCTGCTGATGCTCTTGTCCTCAATTTGGAGCACAGCC     | 600 |
| Qy | 849 | TGGGGCCCTTGGTCACCAACCTGTGGGCTGGGCAT                          | 883 |
| Db | 601 | TGGGGCCCTTGGTCACCAACCTGTGGGCTGGGCAT                          | 635 |

RESULT 4  
AAZ07521  
ID AAZ07521 standard; cDNA; 681 BP.

AAZ07521;

DT 26-NOV-1999 (first entry)

Rat HICP mature polypeptide

KW Heparin-induced CCN-like protein; HICP; cell

|    |   |
|----|---|
| XX | Cardiovascular disorder; aberrant cell proliferation; fibrotic disorder |
| OS | Rattus sp.  |

XX  
PN  
WO9947556-1XX  
PD 23-SEP-1999.

XX  
PF  
18-MAR-1999;

XX  
PR 19-MAR-1998; 98US-0044273.

XX  
PA (TUFT ) TUFTS COLLEGE.XX  
PI  
Castellot JJ;

XX WPI; 1999-562060/47.

DR P-PSDB; AAYZ/44U.  
XX

PT Nucleic acid sequences encoding rat heparin-induced CCN-like protein, used in methods to identify modulators or in diagnostic applications

PS Disclosure; Fig 2; 108pp; English.

The invention provides a rat heparin-induced CGN-like protein (HICP) protein. Agents that stimulate or inhibit HICP protein activity or expression, antisense HICP nucleic acid molecules and HICP antibodies, can be used to modulate cell-associated activity. HICP modulators can be used to treat disorders characterized by aberrant HICP protein activity or expression. Probes capable of hybridizing to HICP mRNA or antibodies specific for HICP can be used to detect HICP activity in a biological sample. HICP can be used to treat disorders, such as a cardiovascular or fibrotic disorder, characterized by aberrant cell proliferation. The present sequence represents the coding sequence of rat HICP mature polypeptide.

Sequence 681 BP; 121 A; 211 C; 215 G; 134 T; 0 other;  
XX

Query Match 64.1%; Score 566; DB 20; Length 681;

Best Local Similarity 100.0%; Pred. No. 4.5e-147;  
Matches 566; Conservative 0; Mismatches 0; Indels

QY 318 CAGCTGTGCCGGACACCCCTGTACCTGTCTTGGACACCACCCCAAGTGGCCACAGGGGGTA 377

Db 1 CAGCTGTCCGGACACCCCTGTACCTGTCTTGGACACCACCCCAAGTGCCCAAGGGGTA 60

QY 378 CCCCTGGTGTGGATGGCTGTGGCTGCTGTAAAGTGTGTGCACGGAGGCTGGGGGAGTCC 437

|     |   |     |
|-----|---|-----|
| 61  | CCCCTGTGTGGATGGCTTGGCTGTCTAAAGTGTGTGCACGGAGGCTGGGGAGTCC     | 120 |
| 438 | TGCACCACTGCATGTCTGGACCCACCGAGGCCCTGGTTTGTACGCTGGGCGAGCG     | 497 |
| 121 | TGCACCACTGCATGTCTGGACCCACCGAGGCCCTGGTTTGTACGCTGGGCGAGCG     | 180 |
| 498 | CCCTGGCGCCATGGGGTGTGTGTCTTTGGATGAGGATGACGGTAGCTGTGAGTGAAT   | 557 |
| 181 | CCCTGGCGCCATGGGGTGTGTGTCTTTGGATGAGGATGACGGTAGCTGTGAGTGAAT   | 240 |
| 558 | GGCCGCAAGTACCTGGATGGAGAGACCTTTAAACCCAAATTCACGGTCTCTGTCGCGCT | 617 |
| 241 | GGCCGCAAGTACCTGGATGGAGAGACCTTTAAACCCAAATTCACGGTCTCTGTCGCGCT | 300 |
| 618 | GATGACGGTGGCTTCACCTGCTGCGCTGTGCAGTGAGGATGTGCGGCTGCCACAGCTGG | 677 |
| 301 | GATGACGGTGGCTTCACCTGCTGCGCTGTGCAGTGAGGATGTGCGGCTGCCACAGCTGG | 360 |
| 678 | GACTGCCACGCGCAGAGAGAAATACAGGTGCCAGAAAGTGTGCCCGCAGTGGGTATGT  | 737 |
| 361 | GACTGCCACGCGCAGAGAGAAATACAGGTGCCAGAAAGTGTGCCCGCAGTGGGTATGT  | 420 |
| 738 | GACCAGGAGTGAACCGGCGATCCAGCGCTCCACGGCGCAGGACACCAACTTTCTGCC   | 797 |
| 421 | GACCAGGAGTGAACCGGCGATCCAGCGCTCCACGGCGCAGGACACCAACTTTCTGCC   | 480 |
| 798 | CTTGCTACTCTGTGCTGTGATGTCTCTTGTCCAAATTTGGAGCACACGCTGGGGCCCC  | 857 |
| 481 | CTTGCTACTCTGTGCTGTGATGTCTCTTGTCCAAATTTGGAGCACACGCTGGGGCCCC  | 540 |
| 858 | TGCTCAACCACTGTGGGCTGGGCAT                                   | 883 |
| 541 | TGCTCAACCACTGTGGGCTGGGCAT                                   | 566 |

## RESULT 5

RESOLUT  
AAX76489/C

XX  
FMA/0405 STAINING, DNA, 733 DE

AC  
XX  
HAA/0405,DI 06-AUG-1999 (FIRST ENTRY)  
XX

DE Mouse WISP-2 protein complementary nucleotide sequence SEQ ID NO:18:

WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;  
connective tissue growth factor; cancer; melanoma; arteriosclerosis;  
leukaemia; lymphoid malignancy; haematopoiesis-related disorder;  
tissue-growth disorder; skin disorder; osteopetrosis; fibrotic lesion;  
kidney disorder; bone-related disorder; osteoporosis; trauma; burn;  
connective tissue disorder; catabolic state; inflammation;  
tissue-related disorder; angiogenesis; immunological disorder; ss

XX Mus. sp.

XX PN WO9921998-A1.

XX PD 06-MAY-1999

XX  
PF  
29-OCT-1998: 98W0-1152'2991

XX  
PR 14-APR-1998: 98US-0081695.

|    |              |               |
|----|--------------|---------------|
| PR | 29-OCT-1997; | 97US-0063704. |
| PR | 03-FEB-1998: | 98US-0073612. |

XX  
PA (GETH ) GENENTECH INC.

PI Botstein DA, Cohen RL, Goddard A, Gurney AL, Hillan K:

PI Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WI, XX

DR WPI; 1999-337420/28.  
XX















|          |     |   |     |
|----------|-----|---|-----|
| QY       | 363 | TGCCACAGGGGTTACCCTGGTGGATGGCTGTGGCTGCCTGTAAGATGTGTGCACGG                        | 422 |
| Db       |     |   |     |
| Db       | 124 | TGCCCGCTGGGAGTACCCTGGTGGATGGCTGTGGCTGCCTGCCGGGTATGTGCACGG                       | 183 |
| QY       | 423 | AGSCTGGGGGAGTCCTCGGACCACTGCATGCTCGCACCCACAGCCAGCGCCTGGTTGT                      | 482 |
| Db       |     |   |     |
| Db       | 184 | CGBCTGGGGGAGCCCTCGGACCAACTCACGCTGCGACGCCACAGCGCCTGGTCTGC                        | 243 |
| QY       | 483 | CAGCCTGGGGCAGGCCCCTGGCGGCATCGGGCTGTGTGCTCTTGATGAGGATGACGCT                      | 542 |
| Db       |     |   |     |
| Db       | 244 | CAGCCGGGGCAGGACCCGGTGGCGGGGGCCCTGTGCTCTTGGCAGAGACGACAGC                         | 303 |
| QY       | 543 | AGCTGTGAGGTGAATPGGCCGAGTACTCTGATGGAGAGACCTTTAAACCCCAATTGCAG                     | 602 |
| Db       |     |   |     |
| Db       | 304 | AGCTGTGAGGTGAACGGCGCCTGTATCGGGAAGGGGAGACCTTCCAGCCCCACTGCAGC                     | 363 |
| QY       | 603 | GTCCTGTGCCCTGTGATGACGGTGGCTTCACTGCGCTGCCGTGTGCAGTAGGATGTG                       | 662 |
| Db       |     |   |     |
| Db       | 364 | ATCCGCTGCCGCTGCGAGGAGCGCGGCTTCACTTGCCTGCCGTGTCCAGCGAGGATGTG                     | 423 |
| QY       | 663 | CGGCTGCCACAGCTGGGACTGCCACGGCCCCAAGAAGATAACAGTGCACGAAAAGTCTGC                    | 722 |
| Db       |     |   |     |
| Db       | 424 | CGGCTGCCACAGCTGGGACTGCCCCACCCACGAGGAGTCCAGTCTTGGCAAGTCTGC                       | 483 |
| QY       | 723 | CCGAGTGGGTATGTGACAGGAGAGTGACACCGCGCATCCAGCGCTCCACGGCGCAAGA                      | 782 |
| Db       |     |   |     |
| Db       | 484 | CCTGAGTGGTGTGGGGCCAAGGAGGGGACTGGGGACCCAGCCCTTCCAGCCCAAGA                        | 543 |
| QY       | 783 | CACCAACTTTTCGCCCTGTGCATCTGCTGCTCTGCTGATGCTCCTTGTCCAAATTGGAGC                    | 842 |
| Db       |     |   |     |
| Db       | 544 | CCCCAGTTTTCTGGCCTTGTCTCTTCCCTGCGCCCTGGTGTGCCCTGCCAGAATGGAGC                     | 603 |
| QY       | 843 | ACAGCCTGGGGCCCTGCTCAACCACTGTGGGCTGGGCAT   | 883 |
| Db       |     |   |     |
| Db       | 604 | ACGGCTGGGACCCCTGCTCGAACCACTGTGGGCTGGGCAT  | 644 |
| RESULT   | 13  |   |     |
| AAA77566 | ID  | AAA77566 standard; cDNA; 1267 BP.   |     |
| XX       | XX  | AAA77566;   |     |
| AC       | XX  |   |     |
| DT       | XX  | 07-NOV-2000 (first entry)   |     |
| DE       | XX  | Human PRO261 cDNA sequence SEQ ID NO:71.  |     |
| KW       | XX  | Human; PRO; promotion; inhibition; angiogenesis; cancer; cardiovascularisation; |     |
| KW       | XX  | diagnosis; trauma; wound; cancer; atherosclerosis; cardiac hypertrophy;         |     |
| KW       | XX  | angiogenic; proliferative; cardiant; cardiovascular; antiatherosclerotic;       |     |
| KW       | XX  | cytostatic; gene therapy; vaccine; ss.  |     |
| OS       | XX  | Homo sapiens.   |     |
| PN       | XX  | WO200032221-A2.   |     |
| PN       | XX  | 08-JUN-2000.  |     |
| PD       | XX  | 30-NOV-1999; 99WO-US28313.  |     |
| PF       | XX  | 01-DEC-1998; 98WO-US25108.  |     |
| PR       | XX  | 16-DEC-1998; 98US-0112850.  |     |
| PR       | XX  | 12-JAN-1999; 98US-0115554.  |     |
| PR       | XX  | 08-MAR-1999; 99WO-US05028.  |     |
| PR       | XX  | 12-MAR-1999; 99US-0123957.  |     |
| PR       | XX  | 28-APR-1999; 99US-0131445.  |     |
| PR       | XX  | 14-MAY-1999; 99US-0134287.  |     |
| PR       | XX  | 02-JUN-1999; 99WO-US12252.  |     |
| PR       | XX  | 23-JUN-1999; 99US-0141037.  |     |
| PR       | XX  | 20-JUL-1999; 99US-0144758.  |     |
| PR       | XX  | 26-JUL-1999; 99US-0145698.  |     |
| PR       | XX  | 01-SEP-1999; 99WO-US20111.  |     |

|    |   |               |
|----|---|---------------|
| PR | 08-SEP-1999;  | 99WO-US20594. |
| PR | 13-SEP-1999;  | 99WO-US20944. |
| PR | 15-SEP-1999;  | 99WO-US21090. |
| PR | 15-SEP-1999;  | 99WO-US21547. |
| PR | 05-OCT-1999;  | 99WO-US23089. |
| PR | 29-OCT-1999;  | 99US-0162506. |
| XX | (GETH ) GENENTECH INC.  |               |
| PA | Ashtenazi AJ, Baker KP, Ferrara N, Gerber H, Hillan KJ, Goddard A;      |               |
| PI | Godowski PJ, Gurney AL, Klein RD, Kuo SS, Paoni NF, Smith V;            |               |
| PI | Watanabe CK, Williams PM, Wood WI;                                      |               |
| XX | WPT; 2000-412154/35.  |               |
| DR | P-PSDB; AAB24402.   |               |
| XX |   |               |
| CC | Nucleic acids encoding PRO polypeptides useful for preventing,          |               |
| PT | diagnosing and treating disorders in mammals -                          |               |
| XX | angiogenic disorders in mammals -                                       |               |
| PS | Claim 61; Fig 29; 315pp; English.                                       |               |
| XX |   |               |
| CC | The present invention describes nucleic acids encoding PRO polypeptides |               |
| CC | useful for preventing, diagnosing and treating disorders in             |               |
| CC | cardiovascular, endothelial or angiogenic disorder in mammals by        |               |
| CC | modulating cell proliferation, angiogenesis and cardiovascularisation,  |               |
| CC | and for identifying agonists and antagonists of these processes. The    |               |
| CC | nucleic acids and the proteins they encode may be used in the           |               |
| CC | prevention, treatment and diagnosis of diseases associated with         |               |
| CC | inappropriate PRO expression such as cardiovascular, endothelial or     |               |
| CC | angiogenic disorders in mammals (e.g. atherosclerosis, cancers and      |               |
| CC | cardiac hypertrophy). For example, the nucleic acids (NCs) and vectors  |               |
| CC | containing them and the PRO polypeptide may be used to treat disorders  |               |
| CC | associated with decreased PRO expression. AAA77510 to AAA77721 and      |               |
| CC | AAB24388 to AAB24435 represent nucleotide and protein sequences used in |               |
| CC | the exemplification of the present invention.                           |               |
| XX |   |               |
| SQ | Sequence 1267 BP; 216 A; 419 C; 390 G; 242 T; 0 other;                  |               |
|    | Query Match 49.2%; Score 434.6; DB 21; Length 1267;                     |               |
|    | Best Local Similarity 79.9%; Pred. No. 1.6e-110;                        |               |
|    | Matches 512; Conservative 0; Mismatches 129; Indels 0; Gaps 0;          |               |
| Qy | 243 GGTGACATGAGGGCAGCCACCATTCTTGGCCACTTCCTTCCTCTGCCCTTC                 | 302           |
| Db |   |               |
|    | 4 GGGACATGAGGCACACCGAAGACCCACTCTTGCCCTTCCTCTGCCCTTC                     | 63            |
| Qy | 303 TCAATGTGTGTGCCAGCTGTGCGGACACCCGTGTACCTGTCTTTGGACACCC                | 362           |
| Db |   |               |
|    | 64 TCAAAGTGGCTACCCAGCTGTGCCGACACCATGTACCTGCCCTGGCACCT                   | 123           |
| Qy | 363 TGCCCCAGGGGTACCCCTGGTGTGGATGGCTGTGGCTGTAAAGTGTGTGACGG               | 422           |
| Db |   |               |
|    | 124 TGCCCCGTGGGAGTACCCCTGGTGTGGATGGCTGTGGCTGTATGTGACCG                  | 183           |
| Qy | 423 AGGCTGGGGAGTCTGCGACACCACTGTCTGCGACCCCGACGAGGCCCTGT                  | 482           |
| Db |   |               |
|    | 184 CGGCTGGGGAGCCCTGCGACAACCTCCAGCTCTGCGACCGCAGGCGCTGT                  | 243           |
| Qy | 483 CAGCCTGGGCGAGGCCCTTGGCGGCCATGGGGCTGTGTGTCTTTGGATGAGG                | 542           |
| Db |   |               |
|    | 244 CAGCCCGGCGAGGACCCCGGTGGCGGGGGCCCTGTGCTCTTTGCGAGGAG                  | 303           |
| Qy | 543 AGCTGTGAGTGAATGGCCCGCAGTACCTGGATGGAGAGACCTTTTAACCCA                 | 602           |
| Db |   |               |
|    | 304 AGCTGTGAGTGAACGGCCGCTGTATCGGGGAAGGGAGACCTTCCAGCCCC                  | 363           |
| Qy | 603 GTCTGTGCGCTGTGATGACGCTTCCACTGCTGCCGCTGTGCACTGAGGATG                 | 662           |
| Db |   |               |
|    | 364 ATCCGTGCCGCTGCGAGGACGGCGCTTCCACTGCTGCCGCTGTGCACTG                   | 423           |
| Qy | 663 CGGCTGCCAGCTGGGACTGCCACGCCCCCAAGAGAATACAGCTGCCAGGAA                 | 722           |
| Db |   |               |







GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OW nucleic - nucleic search, using sw model

Run on: July 28, 2003, 15:55:06 ; Search time 48.04 Seconds  
(without alignments)  
5636.882 Million cell updates/sec

Title: US-10-010-408-l\_COPY\_l\_883

Perfect score: 883

Sequence: 1 GACGCTTCTGATCTCCAGAG.....ACCACCTGTGGCTGGGCAT 883

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued\_Patents\_NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description      |
|------------|-------|-------------|--------|----|------------------|
| 1          | 727.6 | 82.4        | 1734   | 4  | US-09-182-145-17 |
| 2          | 727.6 | 82.4        | 1734   | 4  | US-09-182-145-18 |
| 3          | 434.6 | 49.2        | 1293   | 4  | US-09-182-145-13 |
| 4          | 434.6 | 49.2        | 1293   | 4  | US-09-182-145-14 |
| 5          | 423   | 47.9        | 738    | 4  | US-09-182-145-38 |
| 6          | 420.6 | 47.6        | 841    | 4  | US-09-182-145-39 |
| 7          | 128.4 | 14.5        | 2075   | 1  | US-08-167-628-1  |
| 8          | 128.4 | 14.5        | 2075   | 1  | US-08-386-680-1  |
| 9          | 128.4 | 14.5        | 2075   | 1  | US-08-459-717-1  |
| 10         | 128.4 | 14.5        | 2075   | 1  | US-08-712-302-1  |
| 11         | 128.4 | 14.5        | 2075   | 2  | US-08-880-031-1  |
| 12         | 128.4 | 14.5        | 2075   | 3  | US-09-097-179-1  |
| 13         | 128.4 | 14.5        | 2075   | 4  | US-09-080-715-1  |
| 14         | 128.4 | 14.5        | 2075   | 4  | US-09-142-569-7  |
| 15         | 128.4 | 14.5        | 2075   | 5  | PCT-US96-08140-1 |
| 16         | 128.4 | 14.5        | 2998   | 3  | US-09-054-368-1  |
| 17         | 128.4 | 14.5        | 2998   | 3  | US-09-054-274-1  |
| 18         | 128.4 | 14.5        | 2998   | 4  | US-09-056-704-1  |
| 19         | 125.8 | 14.2        | 2350   | 4  | US-09-187-478-1  |
| 20         | 124.2 | 14.1        | 2350   | 4  | US-09-292-036-1  |
| 21         | 123.4 | 14.0        | 1418   | 4  | US-09-142-569-3  |
| 22         | 120.8 | 13.7        | 2267   | 4  | US-09-142-569-5  |
| 23         | 117   | 13.3        | 1480   | 4  | US-09-142-569-1  |
| 24         | 105.6 | 12.0        | 1766   | 4  | US-09-182-145-9  |
| 25         | 105.6 | 12.0        | 1766   | 4  | US-09-182-145-10 |
| 26         | 102.8 | 11.6        | 1128   | 2  | US-08-459-101A-1 |
| 27         | 97    | 11.0        | 1062   | 4  | US-09-253-316-3  |

|      |      |      |         |   |                   |                   |
|------|------|------|---------|---|-------------------|-------------------|
| 28   | 96.2 | 10.9 | 2830    | 4 | US-09-182-145-1   | Sequence 1, Appl  |
| c 29 | 96.2 | 10.9 | 2830    | 4 | US-09-182-145-2   | Sequence 2, Appl  |
| 30   | 84   | 9.5  | 4214    | 4 | US-09-123-135-1   | Sequence 1, Appl  |
| 31   | 75.4 | 8.5  | 1142    | 4 | US-09-253-316-1   | Sequence 1, Appl  |
| 32   | 75.4 | 8.5  | 1212    | 4 | US-09-182-145-34  | Sequence 34, Appl |
| c 33 | 75.4 | 8.5  | 1212    | 4 | US-09-182-145-35  | Sequence 35, Appl |
| 34   | 75.4 | 8.5  | 1335    | 4 | US-09-182-145-30  | Sequence 30, Appl |
| c 35 | 75.4 | 8.5  | 1335    | 4 | US-09-182-145-31  | Sequence 31, Appl |
| 36   | 73.8 | 8.4  | 1403    | 4 | US-09-182-145-23  | Sequence 23, Appl |
| 37   | 65.8 | 7.5  | 1101    | 4 | US-09-182-145-29  | Sequence 29, Appl |
| 38   | 63.4 | 7.2  | 693     | 4 | US-09-182-145-24  | Sequence 24, Appl |
| 39   | 63.4 | 7.2  | 1202    | 4 | US-09-182-145-26  | Sequence 26, Appl |
| 40   | 48.6 | 5.5  | 2541    | 2 | US-08-656-393-1   | Sequence 1, Appl  |
| 41   | 41.4 | 4.7  | 51      | 4 | US-09-182-145-117 | Sequence 117, App |
| c 42 | 36.6 | 4.1  | 4403765 | 4 | US-09-103-840A-2  | Sequence 2, Appl  |
| c 43 | 36.6 | 4.1  | 4411529 | 4 | US-09-103-840A-1  | Sequence 1, Appl  |
| 44   | 35.8 | 4.1  | 4080    | 2 | US-08-710-249-3   | Sequence 3, Appl  |
| 45   | 35.8 | 4.1  | 4080    | 4 | US-08-220-157A-3  | Sequence 3, Appl  |

#### ALIGNMENTS

##### RESULT 1

US-09-182-145-17  
; Sequence 17, Application US/09182145B

; Patent No. 6387657

; GENERAL INFORMATION:

; APPLICANT: Botstein, David A.

; APPLICANT: Cohen, Robert

; APPLICANT: Goddard, Audrey

; APPLICANT: Gurney, Austin L.

; APPLICANT: Hillan, Kenneth J.

; APPLICANT: Lawrence, David A.

; APPLICANT: Levine, Arnold J.

; APPLICANT: Pennica, Diane

; APPLICANT: Roy, Margaret Ann

; APPLICANT: Wood, William I.

; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME

; FILE REFERENCE: P117682

; CURRENT APPLICATION NUMBER: US/09/182,145B

; CURRENT FILING DATE: 1998-10-29

; EARLIER APPLICATION NUMBER: US 60/063,704

; EARLIER FILING DATE: 1997-10-29

; EARLIER APPLICATION NUMBER: US 60/073,612

; EARLIER FILING DATE: 1998-02-04

; EARLIER APPLICATION NUMBER: US 60/081,695

; EARLIER FILING DATE: 1998-04-14

; NUMBER OF SEQ ID NOS: 156

; SEQ ID NO 17

; LENGTH: 1734

; TYPE: DNA

; ORGANISM: Mus musculus

US-09-182-145-17

Query Match 82.4%; Score 727.6; DB 4; Length 1734;

Best Local Similarity 91.9%; Pred. No. 1.2e-189;

Matches 816; Conservative 0; Mismatches 59; Indels 13; Gaps 4;

Oy 3 CGCTTCTGATCTCCAGAGGACCCCTGGGCTGGGACAGGGGCCCTTGGCAAGGCTGACAGCCGC 62

Db 13 CGCTCCTGATCTCCAGAGGACCCCGGCTGGGACAGGGGCCCTTGGCGAGGCTGACAGTCG 72

Oy 63 TG-GGCAGTGGCTTGGATGGAGTCTTTATCTGGAAGTGGAGCTAGAGGCTCC 121

Db 73 TGTGGCAGTGGCTTGGATGGAGTCTTTTCTTGTGGAGTGGAGGCTGAGAGGCTCC 132

Oy 122 TGTGAG---CTTGTCCTAAAGCTTAGCAGCTTGGTGGCTTGGCTTACACACTGTCA 178

Db 133 TGTGAGGCTCTCTGCTTAAACTCTTGGCACTTGGGCTTGGCTTACACACTGTCA 192

Oy 179 GACACCTTCGTGGTGGGCTCCACGGCCCTCACCTTACGTTTGAAGCTGGCTCCACAAGG 238

Db 179 GACACCTTCGTGGTGGGCTCCACGGCCCTCACCTTACGTTTGAAGCTGGCTCCACAAGG 238

Db 193 GACACCTTCTTGGTGGCCCTCTCTGGCC-----TCAGTTTGAAGCTGGCTCCACAAAGG 246  
QY 239 ACAGGTGACATGAGGGGAGCCCACTGATCATCTTCTGGCCACTTCTTCTCTGCT 298  
Db 247 ACAGGTGACATGAGGGGAGCCCACTGATCATCTTCTGGCCACTTCTTCTCTGCT 306  
QY 299 TCTCTCAATGTTGTGTCCTGAGTGTGTCGAGTGTGTCGAGTGTGTCGAGTGTGTC 358  
Db 307 TCTCTCAATGTTGTATTCAGTGTGTCGAGTGTGTCGAGTGTGTCGAGTGTGTC 366  
QY 359 CCAGTGGCCACAGGGGTACCCCTGCTGATGAGTGTGTCGAGTGTGTCGAGTGTGTC 418  
Db 367 CCAGTGGCCACAGGGGTACCCCTGCTGATGAGTGTGTCGAGTGTGTCGAGTGTGTC 426  
QY 419 ACAGAGGTGGGGAGTCTCTGACACCTGCTGTCGACACCTGCTGTCGACACCTGCT 478  
Db 427 ACAGAGGTGGGGAGTCTCTGACACCTGCTGTCGACACCTGCTGTCGACACCTGCT 486  
QY 479 TTGTCAGCTGGGGAGCCCTGCGGCGCATGAGGCTGTGTCGTCGTCGTCGTCGTCGTC 538  
Db 487 TTGTCAGCTGGGGAGCCCTGCGGCGCATGAGGCTGTGTCGTCGTCGTCGTCGTCGTC 546  
QY 539 CGGTAGCTGTGAGTGAATGCGGCGCATGAGGCTGTGTCGTCGTCGTCGTCGTCGTC 598  
Db 547 CGGTAGCTGTGAGTGAATGCGGCGCATGAGGCTGTGTCGTCGTCGTCGTCGTCGTC 606  
QY 599 CAGGTCTCTGTCGCTGTGATGAGGCTGTGTCGTCGTCGTCGTCGTCGTCGTCGTC 658  
Db 607 CAGGTCTCTGTCGCTGTGATGAGGCTGTGTCGTCGTCGTCGTCGTCGTCGTCGTC 666  
QY 659 TGTGGGCTGCGGAGTGGGCTGCGGCGCATGAGGCTGTGTCGTCGTCGTCGTCGTCGTC 718  
Db 667 TGTGGGCTGCGGAGTGGGCTGCGGCGCATGAGGCTGTGTCGTCGTCGTCGTCGTCGTC 726  
QY 719 CTGCCCCAGTGGGTATGTACAGGAGTGA---CACCGGGATCCAGCGCTCCACGCG 775  
Db 727 CTGCCCCAGTGGGTATGTACAGGAGTGA---CACCGGGATCCAGCGCTCCACGCG 786  
QY 776 GCAAGGACACCACTTTCTGCGCTGTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 835  
Db 787 GCAAGGACACCACTTTCTGCGCTGTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 846  
QY 836 TTGAGGACACCTGGGGCCCTGCTCAACCACTGTGGGCTGGGCAT 883  
Db 847 CTGGAGCACAGCCTGGGGCCCTGCTCAACCACTGTGGGCTGGGCAT 894

## RESULT 2

US-09-182-145-18/c  
; Sequence 18, Application US/09182145B  
; Patent No. 6387657  
; GENERAL INFORMATION:  
; APPLICANT: Botstein, David A.  
; APPLICANT: Cohen, Robert  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth J.  
; APPLICANT: Lawrence, David A.  
; APPLICANT: Levine, Arnold J.  
; APPLICANT: Pennica, Diane  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Wood, William I.  
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: P1176R2  
; CURRENT APPLICATION NUMBER: US/09/182,145B  
; CURRENT FILING DATE: 1998-10-29  
; EARLIER APPLICATION NUMBER: US 60/063,704  
; EARLIER FILING DATE: 1997-10-29  
; EARLIER APPLICATION NUMBER: US 60/073,612  
; EARLIER FILING DATE: 1998-02-04  
; EARLIER APPLICATION NUMBER: US 60/081,695  
; EARLIER FILING DATE: 1998-04-14  
; NUMBER OF SEQ ID NOS: 156

; SEQ ID NO 18  
; LENGTH: 1734  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-09-182-145-18

Query Match 82.4%; Score 727.6; DB 4; Length 1734;  
Best Local Similarity 91.9%; Pred. No. 1.2e-189;  
Matches 816; Conservative 0; Mismatches 59; Indels 13; Gaps 4;

QY 3 GCCTTCTGATCTCAGAGGACCCCTGGGTGGGACAGGGCCCTTGGCAAGGCTGCAGCCGC 52  
Db 1722 GCCTTCTGATCTCAGAGGACCCCGGCTGGGACAGGGCCCTTGGCGAGGCTGCAGCTGC 1663  
QY 63 TG-GCAGTGGCTTGGAAATGAGAGTCTTTATTACTGGGAACCTGAGAGCTAAGAGCTCC 121  
Db 1662 TGTGGCAGTAGCTGGGATGAGAGTCTTTCTTCTGCTGGNACTGAGAGCTGAGAGCTCC 1603  
QY 122 TGTAG---CTTGTCTTAAAGTCTTAGCACTTTGTGGTGGCTTGTACACACTGTCA 178  
Db 1602 TGTACAGGCTCTGTGCTTAAACTCTTGGCACTTGGCGTGGCTTGGGCTTCCACACACTGTCA 1543  
QY 179 GACACCTTCTGGTGGCTTCCAGGCGCTCAGCTTTCAGGTTTGAAGCTGCCTCCACAGGG 238  
Db 1542 GACACCTTCTGGTGGCTTCTCTCGGCC-----TCAGGTTTGAAGCTGGCTCCACAGGG 1489  
QY 239 ACAGGTGACATGAGGGGAGCCCACTGATCATCTTCTGGCCACTTCTTCTCTGCTGCT 298  
Db 1488 ACAGGTGACATGAGGGGAGCCCACTGATCATCTTCTGGCCACTTCTTCTCTGCTGCT 1429  
QY 299 TCTCTCAATGTTGTGTCCTGAGTGTGTCGCGGACACCTGTACCTTGTCTTGTGACACCTCC 358  
Db 1428 TCTCTCAATGTTGTATTCAGCTGTGCGGAGCCCTGTGCTGTGCTTGTGACACCTCC 1369  
QY 359 CCAGTGGCCACAGGGGTACCCCTGCTGCTGATGAGTGTGCTGCTGCTGCTGCTGCTGCT 418  
Db 1368 CCAGTGGCCACAGGGGTACCCCTGCTGCTGATGAGTGTGCTGCTGCTGCTGCTGCTGCT 1309  
QY 419 ACAGAGGTGGGGAGTCTCTGCGACACCTGCTGCTGCGACCTGCTGCTGCTGCTGCTGCTGCT 478  
Db 1308 ACAGAGGTGGGGAGTCTCTGCGACACCTGCTGCTGCGACCTGCTGCTGCTGCTGCTGCTGCT 1249  
QY 479 TTGTCAGCCTTGGGAGGCTTGGGCGCATGGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 538  
Db 1248 TTGTCAGCCTTGGGAGGCTTGGGCGCATGGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1189  
QY 539 CGGTAGCTGTGAGTGAATGCGCGAGGTACCTGATGAGAGACCTTTAAACCCCAATTG 598  
Db 1188 CGGAGCTGTGAGTGAATGCGCGAGGTACCTGATGAGAGACCTTTAAACCCCAATTG 1129  
QY 599 CAGGTCTCTGCTGCTGCTGATGAGGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 658  
Db 1128 CAGGTCTCTGCTGCTGATGAGGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1069  
QY 659 TGTGGGCTGCGGAGCTGGGCTGCGGCGCATGGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 718  
Db 1068 TGTGGGCTGCGGAGCTGGGCTGCGGCGCATGGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1009  
QY 719 CTGCCCCAGTGGGTATGTACAGGAGTGA---CACCGGGATCCAGCGCTCCACGCG 775  
Db 1008 CTGCCCCAGTGGGTATGTACAGGAGTGA---CACCGGGATCCAGCGCTCCACGCG 949  
QY 776 GCAAGGACACCACTTTCTGCGCTTGTCACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 835  
Db 948 CCAAGGACACCACTTTCTGCGCTTGTCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 889  
QY 836 TTGAGGACACCTTGGGCGCTTGTCAACCACTGTGGGCTGGGCAT 883  
Db 888 CTGGAGCACAGCCTGGGCGCTTGTCAACCACTGTGGGCTGGGCAT 841

RESULT 3  
US-09-182-145-13











Db 195 GCGCGCCCTGCGCCAGAACTACAGCGGCGCTGCGGTCGCCGAGCAGCGCGCGCG 254  
QY 362 GTGCCACAGGGGTACCCCTGGTCTGTGATGGCTGTGGCTGTAAAGTGTGTCAGC 421  
Db 255 CTGCCCGCGGGCTGAGCTGTGCTGAGCGGTGCGGCTGTCCCGCTCTGCGCCAA 314  
QY 422 GAGGCTGGGGAGTCTCGCACCACTGCATGTCTGCGACCCCGCAGGCGCTGTGTTG 481  
Db 315 GCAGCTGGCGAGCTGTGACCGAGCGAGCCCTGCGACCGCACAAGGCGCTCTTCTG 374  
QY 482 TCAGCTGGCGAGCGCTGCGGCGCATGCGGCTGTGCTCTTGGATGAGGATGCGG 541  
Db 375 TGACTTGGCTCCCGCGCAACCGCAAGATCGGCGTGTG---CACCGCAAGATGTTGC 431  
QY 542 TAGCTGTGAGTGAATGCGCGCAGGTACCTGGATGAGAGACCTTTAAACCAATTCGAG 601  
Db 432 TCCCTGCATCTTCGGTGTACCGCAGCGGAGAGTCTTCCAGAGCAGCTGCAA 491  
QY 602 GGTCTGTGCGCTGTGATGAGCGGTTCACCTGCGCTGCGCTGTGTCAGTGAAGTGT 661  
Db 492 GTACCACTGTCAGTGTGAGCGGCGGTGGGCTGCATGCGCTGTGTCAGATGGAGT 551  
QY 662 GCGGCTGCGGAGTGGGAGTGGCCAGCGCCCAAGAGATACAGTGCAGGAAAGTGTG 721  
Db 552 TCGTCTGCGGAGCTGACCTGCGCTTCCCGAGGAGGTCAAGCTGCGCGGAAATGCTG 611  
QY 722 CCGGAGTGGGTATGTACCGAGGAGTGACACCGCGCATCCAGCGCTCCACGGCGCAAG 781  
Db 612 CGAGAGTGGGTGTGACGAGCCCAAGGACCAACCGTGTGGGCTGCGCTCGCGCG 671  
QY 782 ACACCACTTCTGCGCTGTGACCTTCCCTGTGCTGCTGCTGCTGCTGCTGCTGCT 835  
Db 672 TTACCGACTGGAAGACAGCTTTGGCCAGAGCCCACTATGATTAGAGCAACTGCTG 731  
QY 836 TTGGAGCACAGCTGGGCGCCCTGCTCAACCACTGTGGGCTGGGCAT 883  
Db 732 CCAGACCACAGATGGAGCGCTGTTCAGAACCTGTGGGATGGGCAT 779

RESULT 10

US-08-712-302-1  
; Sequence 1, Application US/08712302  
; Patent No. 5783187  
; GENERAL INFORMATION:  
; APPLICANT: Grotendorst, Gary R.  
; APPLICANT: Bradham Jr., Douglas M.,  
; TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Spensley Horn Jubas & Lubitz  
; STREET: 4225 Executive Square, Suite 1400  
; CITY: La Jolla  
; STATE: CA  
; COUNTRY: US  
; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/712.302  
; FILING DATE: 11-SEP-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/386,680  
; FILING DATE: 10-FEB-1995  
; APPLICATION NUMBER: US/08/167,628  
; FILING DATE:  
; APPLICATION NUMBER: US/07/752,427  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Wetherell, Jr. Ph.D., John W.

; REGISTRATION NUMBER: 31,678  
; REFERENCE/DOCKET NUMBER: PD-1294  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619-455-5100  
; TELEFAX: 619-455-5110  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2075 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; IMMEDIATE SOURCE:  
; CLONE: DB60R32  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 130..1177  
US-08-712-302-1  
  
Query Match 14.5%; Score 128.4; DB 1; Length 2075;  
Best Local Similarity 52.2%; Pred. No. 7.7e-26;  
Matches 338; Conservative 0; Mismatches 301; Indels 9; Gaps 2;  
  
QY 242 CGGTGACATGAGGGCAGCCCACTGATCCATCTTCTTGCCACTTCTTCTCTCTCTCT 301  
Db 135 GCGCGCCAGTATGGCGCCCGTCCGCTGCTGCTGCTCTCTCTCTCTCTCTCTCTCTCG 194  
QY 302 CTCAATGTGTGTGCCAGCTGTCCGCGACACCTGTACCTGTCTCTTGACACACCCCA 361  
Db 195 GCGCGCCCTGCGCGAGAACTGCAGCGGCGCTGCGGTCGCCGAGCAGCGCGCGCG 254  
QY 362 GTGCCACAGGGGTACCCCTGGTGTGCTGATGGCTGTGGCTGTCTGTAAGTGTGTGTCAG 421  
Db 255 CTGCCCGCGCGCTGAGCCTCTGCTGCGCGCTGCGGCTGCTGCGCGCTCTGCGCCAA 314  
QY 422 GAGGCTGGGGAGTCTCGACACCTGCATGTCTGCGACCCCGCAGGCGCTGCTGTTG 481  
Db 315 GCAGCTGGCGAGCTGTGCAGCGCGAGCCCTGCGCAGCCCGCAGCAGGCGCTCTTCTG 374  
QY 482 TCAGCTGGCGAGCGCTGCGGCGCATGCGGCTGTGTCTCTTGGATGAGGATGAGCG 541  
Db 375 TGACTTGGCTCCCGCGCAACCGCAAGATCGGCTGTG---CACCGCAAGATGTTGC 431  
QY 542 TAGCTGTGAGTGAATGCGCGCAGGTACCTGGATGAGAGACCTTTAAACCAATTCGAG 601  
Db 432 TCCCTGCATCTTCGGTGTACCGCAGCGGAGAGTCTTCCAGAGCAGCTGCAA 491  
QY 602 GGTCTGTGCGCTGTGATGAGCGGTTCACCTGCGCTGCGCTGTGTCAGTGAAGTGT 661  
Db 492 GTACCACTGTCAGTGTGAGCGGCGGTGGGCTGCATGCGCTGTGTCAGATGGAGT 551  
QY 662 GCGGCTGCGGAGTGGGACTGCCACAGCGCCCAAGAGATACAGTGCAGGAAAGTGTG 721  
Db 552 TCGTCTGCGGAGCTGACTGCGCTTCCCGAGGAGGTCAAGCTGCGCGGAAATGCTG 611  
QY 722 CCGGAGTGGGTATGTGACCGAGGAGTGACACCGCGCATCCAGCGCTCCAGCGCGCAAG 781  
Db 612 CGAGAGTGGGTGTGACGAGCCCAAGGACCAACCGTGTGGGCTGCGCTCGCGCG 671  
QY 782 ACACCACTTCTGCGCTGTGACCTTCCCTGTGCTGCTGCTGCTGCTGCTGCTGCT 835  
Db 672 TTACCGACTGGAAGACAGCTTTGGCCAGAGCCCACTATGATTAGAGCAACTGCTG 731  
QY 836 TTGGAGCACAGCTGGGCGCCCTGCTCAACCACTGTGGGCTGGGCAT 883  
Db 732 CCAGACCACAGATGGAGCGCTGTTCAGAACCTGTGGGATGGGCAT 779

RESULT 11

US-08-880-031-1  
; Sequence 1, Application US/08880031  
; Patent No. 5916756  
; GENERAL INFORMATION:

APPLICANT: Grotendorst, Gary R.  
APPLICANT: Bradham Jr., Douglas M.,  
TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Spensley Horn Jubas & Lubitz  
STREET: 4225 Executive Square, Suite 1400  
CITY: La Jolla  
STATE: CA  
COUNTRY: US  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/880,031  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/167,628  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Wetherell, Jr. Ph.D., John W.  
REGISTRATION NUMBER: 31,678  
REFERENCE/DOCKET NUMBER: PD-1294  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-455-5100  
TELEFAX: 619-455-5110  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2075 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
IMMEDIATE SOURCE:  
CLONE: DB60R32  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 130...1177  
US-08-880-031-1

Query Match 14.5%; Score 128.4; DB 2; Length 2075;  
Best Local Similarity 52.2%; Pred. No. 7.7e-26;  
Matches 338; Conservative 0; Mismatches 301; Indels 9; Gaps 2;  
QY 242 CGGTGACATGAGGGGCGACCCCACTGATCCATCTTCTGGCCACTTCTTCTCTCTCTCTCT 301  
DB 135 CGCGCCAGATGGGCGCGCGTCCGCGTCTGCTGCTCTCTCTCTCTCTCTCTCTGCGCG 194  
QY 302 CTCATGTGTGTGCGCCAGCTGTGCGGACACCCCTGTACCTGTCTTGGACACCAACCCCA 361  
DB 195 GCCGCGCTGCGCCAGACTGCGAGCGCGCGTGGCGGAGAGCGCGCGCGCG 254  
QY 362 GTGCCACAGGGGTPACCCCTGTGGTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 421  
DB 255 CTGCCCGCGCGGTGACCTCGTGTGACGGCTGCGGCTGCTGCGCGGTCTGCGCGCA 314  
QY 422 GAGGTGGGGAGTCTCGGACCACTGATGTCGACCCCGACCGGAGCGGCGTGGTTG 481  
DB 315 CGAGCTGGGCGAGTGTGACCGGCGGACCCCTGCGACCGCGGACCGGCGGCGCTCTCTG 374  
QY 482 TCAGCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 541  
DB 375 TGACTTGGCTTCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 431  
QY 542 TAGCTGTGAGGTGAATGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 601  
DB 432 TCCCTGCATCTTCGGTGTGTACGGTGTACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 491  
QY 602 GGTCTGTGCGGCGTGTGATGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 661

DB 492 GTACCAGTGCAGTGGCTGGACGGGCGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 551  
QY 662 GCGGCTGCCAGTGGGACTGCCACGCGCCCAAGAGAATACAGGTGCCAGGAAAGTGTG 721  
DB 552 TCGTCTGCCAGCGCTGACTGCCCTTCCGAGGAGGTCAAGCTGCCGCGGAATGTG 611  
QY 722 CCCCAGTGGGTAGTGTGACGAGGAGTGACACCGCGGATCCAGCGCTCCACGCGCAAG 781  
DB 612 CGAGGAGTGGGTGTGTGACGAGCGCCCAAGACCAACCGGTGGTGGCGCTGCCCTCGCGC 671  
QY 782 ACACCAACTTCTGCCCTTGTACCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 835  
DB 672 TTACCGACTGGAAGACACGCTTGGCGCCAGACCACTATGATTAGAGCCAACTGCTGCT 731  
QY 836 TTGAGACACGCTGGGCGCCCTCTCAACCACTGTGGGCTGGGCGAT 883  
DB 732 CCAGACCACAGAGTGGAGCGGCGCTGTTCAGACCTGTGGGATGGCGAT 779

RESULT 12  
US-09-097-179-1  
Sequence 1, Application US/09097179  
Patent No. 6149916  
GENERAL INFORMATION:  
APPLICANT: Grotendorst, Gary R.  
APPLICANT: Bradham Jr., Douglas M.,  
TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Spensley Horn Jubas & Lubitz  
STREET: 4225 Executive Square, Suite 1400  
CITY: La Jolla  
STATE: CA  
COUNTRY: US  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/097,179  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/386,680  
FILING DATE: 10-FEB-1995  
APPLICATION NUMBER: US/08/167,628  
FILING DATE:  
APPLICATION NUMBER: US/07/752,427  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Wetherell, Jr. Ph.D., John W.  
REGISTRATION NUMBER: 31,678  
REFERENCE/DOCKET NUMBER: PD-1294  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-455-5100  
TELEFAX: 619-455-5110  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2075 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
IMMEDIATE SOURCE:  
CLONE: DB60R32  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 130...1177  
US-09-097-179-1

```

Query Match          14.5%; Score 128.4; DB 3; Length 2075;
Best Local Similarity 52.2%; Pred. No. 7.7e-26;
Matches 338; Conservative 0; Mismatches 301; Indels 9; Gaps 2;

QY      242 CGGTGACATGAGGGCAGCCCACTGATCCATCTTCCTGGCCACATTCCTTCCTCGCCTTCT 301
DB      135 CGCGCCAGTAGTGGGCCCGTGCGCGGTGCGCTTCGTGGTCTCCTCGCCCTCTGCACCGG 194

QY      302 CTCAAATGGTGTGTGCCAGAGTGTGTGCCGGACACACCTGTACTGTCTTTGGACACACCACCCA 361
DB      195 GC CGGCCGTGCGCCAGAACACTGCAGCGGGCGGTGCCGGTGC CCGSAGACGCGGGCGCGG 254

QY      362 GTGCCACAGAGGGGTACCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 421
DB      255 CTGCCCGGGGGCGGTGAGCGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 314

QY      422 GAGGCTGGGGAGTCTCTGCAGACCACTGCACTGTCTGTGCGACCCACGACGACGACGACGACGACG 481
DB      315 GCAGCTGGGGAGTCTGTGTGACCGAGCGGACCCCTGCGACCGGACGAAGGGCTCTTTCTGT 374

QY      482 TCAGCTTGGGGCAGGGCCCTGGCGGCCCATGGGGCTGTGTGTCTCTTGGATGAGGATGACGG 541
DB      375 TGACTTTCGGCTCCCCGGCCAAACCGCAAGATCGGGGTGTG ---CACCGCCAAAGATGTGTC 431

QY      542 TAGCTGTGAGGTGAATGGCGCGCAGGTACTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 601
DB      432 TCCCTGCATCTTCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 491

QY      602 GTTCCTGTGCCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 661
DB      492 GTACCAAGTGCAGCTGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 551

QY      662 GCGGCTGCCAGCTGGGAGCTGCCAGCGGCCCAAAGAGAATAACAGGTGCGCAGGAAGTGTGTGTGT 721
DB      552 TGTGTGTGCCAGCGCTGACTGTGCCCTTCCCGAGGAGGGTCAAGTGTGCCCGGGAATGCTGTGTGT 611

QY      722 CCCCGAGTGGGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 781
DB      612 CGAGGAGTGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 671

QY      782 ACACCAACTTTCGCCCTTGTCTACTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 835
DB      672 TTACCGAGCTGGAAGACACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 731

QY      836 TTGGCAGCAGCGCTGGGGCGCCCTGCTCAACACCACTGTGGGTGTGGGCAT 883
DB      732 CCAGACCACAGATGGAGCGCCCTGTTCACAGACCTGTGGGATGGGCAT 779


RESULT 13
US-09-080-715-1
; Sequence 1, Application US/09080715
; Patent No. 6150884
; GENERAL INFORMATION:
; APPLICANT: Grotendorst, Gary R.
; APPLICANT: Bradham Jr., Douglas M.,
; TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spensley Horn Jubas & Lubitz
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/080,715
; FILING DATE:

```

[illegible]

Db 732 CCAGACCAGAGTGGAGCGGCTGTTCACAGACCTGTGGGATGGGCAT 779

RESULT 14  
US-09-142-569-7  
; Sequence 7, Application US/09142569  
; Patent No. 6413735  
; GENERAL INFORMATION:  
; APPLICANT: Lau, Lester F.  
; TITLE OF INVENTION: Extracellular Matrix Signalling Molecules  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States of America  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/142,569  
; FILING DATE: 02-Apr-1999  
; CLASSIFICATION: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Clough, David W.  
; REGISTRATION NUMBER: 36,107  
; REFERENCE/DOCKET NUMBER: 28758/33766  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/474-6300  
; TELEFAX: 312/474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2075 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: "CTGF cDNA coding sequence"  
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:  
US-09-142-569-7

Query Match 14.5%; Score 128.4; DB 4; Length 2075;  
Best Local Similarity 52.2%; Pred. No. 7.7e-26;  
Matches 338; Conservative 0; Mismatches 301; Indels 9; Gaps 2;

QY 242 CGGTGACATGAGGGGAGCCACCTGATCCATCTTGTGCCACTTCTTCTGCTCTCT 301  
DB 135 CGCGCCAGTATGGCCCGCTCCGGTGCCTTGTGTCTCTCTGCTCTCTCTGAGCGG 194  
QY 302 CTCAATGTGTGTGCCAGCTGTGTGCCGACACCTGTACCTGTCTTGGACACCAACCCA 361  
DB 195 GCCGCGCTCGCCAGAACTCCAGGGGCGGTGCCGGTCCCGGACGAGCGCGCGCG 254  
QY 362 GTGCCACAGGGGTACCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 421  
DB 255 CTGCCCCGGGCGGTGAGCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 314  
QY 422 GAGCTGGGGAGTCTCGACCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 481  
DB 315 CGAGCTGGCGGAGT 374  
QY 482 TCAGCTGGGGAGCCCTGTGCCCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 541  
DB 375 TGACTTGGCTCCCGGCCAACCCAGATCGGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 431  
QY 542 TAGCTGTGAGTGAATGGCCCGCAGGTACCTGGATGGAGAGACCTTTAAACCCCAATTGCAG 601

Db 432 TCCCTGCATCTTGGT 491  
QY 602 GGTCTGTGCGCGT 661  
DB 492 GTACCAAGTGCACGTGCTGT 551  
QY 662 GCGCTGCCAGCTGGGACTGCCACGCGCCCAAGAGAAATACAGGTGCCAGGAAAGTGTGTGT 721  
DB 552 TCGTGTGCCAGCTGCTGT 611  
QY 722 CCCGAGTGGGT 781  
DB 612 CGAGGAGTGGGT 671  
QY 782 ACACCAACTTCTGCTGT 835  
DB 672 TTACCGACTGGAAGACACGTTTGGCCAGACCAACTATGATGTAGAGCCAACTGCTGTGT 731  
QY 836 TTGAGACACAGCTGGGCGCTGT 883  
DB 732 CCAGACCACAGAGTGGAGCGGCTGTTCACAGACCTGTGGGATGGGCAT 779

RESULT 15  
PCT-US96-08140-1  
; Sequence 1, Application PC/TUS9608140  
; GENERAL INFORMATION:  
; APPLICANT: University of South Florida  
; TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: FISH & RICHARDSON P.C.  
; STREET: 4225 Executive Square, Suite 1400  
; CITY: La Jolla  
; STATE: CA  
; COUNTRY: US  
; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US96/08140  
; FILING DATE: 30-MAY-1996  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Haile, Ph.D., Lisa A.  
; REGISTRATION NUMBER: 38,347  
; REFERENCE/DOCKET NUMBER: 07414/003W01  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619-678-5070  
; TELEFAX: 619-678-5099  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2075 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; IMMEDIATE SOURCE:  
; CLONE: CTGF  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 130..1177  
PCT-US96-08140-1

Query Match 14.5%; Score 128.4; DB 5; Length 2075;  
Best Local Similarity 52.2%; Pred. No. 7.7e-26;  
Matches 338; Conservative 0; Mismatches 301; Indels 9; Gaps 2;

QY 242 CGGTGACATGAGGGGAGCCACCTGATCCATCTTGTGCCACTTCTTCTGCTCTCT 301  
DB 135 CGCGCCAGTATGGCCCGCTCCGGTGCCTTGTGTCTCTCTGCTCTCTCTGAGCGG 194  
QY 302 CTCAATGTGTGTGCCAGCTGTGTGCCGACACCTGTACCTGTCTTGGACACCAACCCA 361  
DB 195 GCCGCGCTCGCCAGAACTCCAGGGGCGGTGCCGGTCCCGGACGAGCGCGCGCG 254  
QY 362 GTGCCACAGGGGTACCCCTGT 421  
DB 255 CTGCCCCGGGCGGTGAGCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 314  
QY 422 GAGCTGGGGAGTCTCGACCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 481  
DB 315 CGAGCTGGCGGAGT 374  
QY 482 TCAGCTGGGGAGCCCTGTGCCCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 541  
DB 375 TGACTTGGCTCCCGGCCAACCCAGATCGGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 431  
QY 542 TAGCTGTGAGTGAATGGCCCGCAGGTACCTGGATGGAGAGACCTTTAAACCCCAATTGCAG 601



Db 135 CGCGCCAGTATGGCCCGCTCCGCGTGCCTTCGTGTCCTCTCGCCCTCTGACCGG 194  
Qy 302 CTCAATGGTGTGTCGCCAGCTGTGCCGACACACCTGTACCTGTCTTTGGACACACCCCA 361  
Db 195 GCCGGCCGTGGCCAGACTGCACGCGCGCGTGCCTGCCGACGAGCGCGCGCGG 254  
Qy 362 GTGCCCCACAGGGGTACCCCTGTGCTGGATGGCTGTGGCTGTAAAGTGTGACG 421  
Db 255 CTGCCCGCGGGCTGAGCTGTGCTGGACGGTGGCTGTGCGCGCTCTCGCCAA 314  
Qy 422 GAGCTGGGGAGTCTTCGACACACCTGCATGTCTGCGACCCCGACGAGCGCTGTTG 481  
Db 315 GCAGCTGGGGAGTGTGACCGGAGCGCCCTGCGACCCCGACGAGCGCTCTTCTG 374  
Qy 482 TCAGCCTGGGGCAGCGCCCTGGCGCCATGGGGCTGTGTCTCTTGGATGAGGATGACGG 541  
Db 375 TGACTTGGCTCCCGGCCAACCCAGATCGGCGTGTG---CACCCCAAGATGGTGC 431  
Qy 542 TAGCTGTAGGTGAATGGCCGACAGTACCTGGATGGAGAGACCTTTAAACCCAAATGCAG 601  
Db 432 TCCCTGCATCTTCGGTGTACGGTGTACCGGACGGAGAGTCTTCCAGAGCAGCTGCAA 491  
Qy 602 GGTCTGTGCCGTGTGATGACGGTGGCTTCACCTGCCCTGCCGTGTGCAGTGAGGATGT 661  
Db 492 GTACCAGTGCACGTGCTGGACGGGGCGGTGGGTGCTGATGCCCTGTGCAGCATGGACGT 551  
Qy 662 GCGGTGCCCGAGCTGGGACTGCCACGCCCCAAGAGAATACAGGTGCCAGGAAAGTGTG 721  
Db 552 TCGTCTGCCAGCCCTGACTGCCCTTCCCGAGGAGGTCAAGCTGCCCGGGAATGCTG 611  
Qy 722 CCCGAGTGGGTATGTGACACGAGGAGTGACACCGCGCATCCAGCGCTCCACGCGCAAGG 781  
Db 612 CGAGGAGTGGGTGTGTGACGAGCCCAAGGACCAACCGTGGTGGGCTGCCCTCGCGGC 671  
Qy 782 ACACCAACTTCTGCCCTTGTCACTCTGCTCTGC-----TGATGCTCTTGTCCAAA 835  
Db 672 TTACCGACTGGAAGACACAGTTTGGCCCGACGCCAACTATGATTAGACCAACTGCCTGGT 731  
Qy 836 TTGAGCACAGCCTGGGGCCCTGCTCAACACCTGTGGGCTGGGCAT 883  
Db 732 CCAGACCACAGTGGAGCGGCTGTTCCAGACCTGTGGGATGGGCAT 779

Search completed: July 28, 2003, 21:35:54  
Job time : 52.24 secs





LENGTH: 1708 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 249..1001  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-10-010-408-1

Query Match 100.0%; Score 883; DB 15; Length 1708;  
Best Local Similarity 100.0%; Pred. No. 2.2e-248;  
Matches 883; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGCTCTGATCTCCAGAGACCTGGGGTGGGACAGGGCCCTTGGCAAGGCTGCACCC 60  
DB 1 GACGCTCTGATCTCCAGAGACCTGGGGTGGGACAGGGCCCTTGGCAAGGCTGCACCC 60

QY 61 GCTGGCAGTGGCTTGGGAATGAGGCTCTTTATTACTGGAACTGAGGAGCTAAAGGCTC 120  
DB 61 GCTGGCAGTGGCTTGGGAATGAGGCTCTTTATTACTGGAACTGAGGAGCTAAAGGCTC 120

QY 121 CTGTCAGCTTGTCTTAAGTCTTAGCACTTGTGGTGGCTTGGGCTTCACACTGTGAGA 180  
DB 121 CTGTCAGCTTGTCTTAAGTCTTAGCACTTGTGGTGGCTTGGGCTTCACACTGTGAGA 180

QY 181 CACCTTCTGTGGCTCCAGGCTCACCTTCACTTCACTTCACTTCACTTCACTTCACTTCA 240  
DB 181 CACCTTCTGTGGCTCCAGGCTCACCTTCACTTCACTTCACTTCACTTCACTTCACTTCA 240

QY 241 ACGGTGACATGAGGGGACGCCACTGATCCATCTTCTGGCCACTTCTTCTCTGCTTC 300  
DB 241 ACGGTGACATGAGGGGACGCCACTGATCCATCTTCTGGCCACTTCTTCTCTGCTTC 300

QY 301 TCTCAATGGTGTGCCAGCTGTGCCGACACCTGTACCTGTCTTGGACACACCC 360  
DB 301 TCTCAATGGTGTGCCAGCTGTGCCGACACCTGTACCTGTCTTGGACACACCC 360

QY 361 AGTGCCACAGGGGTACCTTGTGCTGGATGGCTGGCTGTGATGATGATGATGATGATG 420  
DB 361 AGTGCCACAGGGGTACCTTGTGCTGGATGGCTGGCTGTGATGATGATGATGATGATG 420

QY 421 GGAGCTGGGGAGTCTCGACACACCTGCATGTGCGACCCACCGACCGAGGCTGTGTT 480  
DB 421 GGAGCTGGGGAGTCTCGACACACCTGCATGTGCGACCCACCGAGGCTGTGTT 480

QY 481 GTCAGCTGGGACAGGCTGGGCGCATGGGCTGTGCTCTTGGATGAGGATGAG 540  
DB 481 GTCAGCTGGGACAGGCTGGGCGCATGGGCTGTGCTCTTGGATGAGGATGAG 540

QY 541 GTAGCTGTGAGTGAATGGCGCAGGTACCTGGATGGAGACCTTTAAACCCAATTGCA 600  
DB 541 GTAGCTGTGAGTGAATGGCGCAGGTACCTGGATGGAGACCTTTAAACCCAATTGCA 600

QY 601 GGGTCTGTGCGCTGTGATGACGCTGGCTTACCTGCTGCGCTGTGCGCTGTGCGCTGTG 660  
DB 601 GGGTCTGTGCGCTGTGATGACGCTGGCTTACCTGCTGCGCTGTGCGCTGTGCGCTGTG 660

QY 661 TGGGCTGCGCTGTGATGACGCTGGCTTACCTGCTGCGCTGTGCGCTGTGCGCTGTG 720  
DB 661 TGGGCTGCGCTGTGATGACGCTGGCTTACCTGCTGCGCTGTGCGCTGTGCGCTGTG 720

QY 721 GCCCGAGTGGTGTGTACAGGAGTGTACACCGGCTGATCCAGGCTTCCAGCGCAAG 780  
DB 721 GCCCGAGTGGTGTGTACAGGAGTGTACACCGGCTGATCCAGGCTTCCAGCGCGCAAG 780

QY 781 GACACCAACTTCTGCCCTTGTCACTCTGCTCTGCTGTGCTGTGCTGTGCTGTGCTGTG 840  
DB 781 GACACCAACTTCTGCCCTTGTCACTCTGCTCTGCTGTGCTGTGCTGTGCTGTGCTGTG 840

QY 841 GCACAGCTGGGGCCCTTGTCAACCACTTGTGGCTGGGAT 883  
DB 841 GCACAGCTGGGGCCCTTGTCAACCACTTGTGGCTGGGAT

DB 841 GCACAGCTGGGGCCCTTGTCAACCACTTGTGGCTGGGAT 883

RESULT 2  
US-10-112-267-17  
; Sequence 17, Application US/10112267  
; Publication No. US2003006878A1  
; GENERAL INFORMATION:  
; APPLICANT: Botstein, David A.  
; APPLICANT: Cohen, Robert  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth J.  
; APPLICANT: Lawrence, David A.  
; APPLICANT: Levine, Arnold J.  
; APPLICANT: Pennica, Diane  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Wood, William I.  
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: P1176R2  
; CURRENT APPLICATION NUMBER: US/10/112,267  
; CURRENT FILING DATE: 2002-03-27  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/182,145B  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-29  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/063,704  
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-10-29  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/073,612  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-04  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/081,695  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-14  
; NUMBER OF SEQ ID NOS: 156  
; SEQ ID NO 17  
; LENGTH: 1734  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-10-112-267-17

Query Match 82.4%; Score 727.6; DB 15; Length 1734;  
Best Local Similarity 91.9%; Pred. No. 6.2e-203;  
Matches 816; Conservative 0; Mismatches 59; Indels 13; Gaps 4;

QY 3 CGCTTCTGATCTCCAGAGACCTTGGGTGGGACAGGGCCCTTGGCAAGGCTGCACGCCG 62  
DB 13 CGCTCTGATCTCCAGAGAGCCCGGCTGGGACAGGGCCCTTGGCGAGGCTGCACGTGC 72

QY 63 TG-GGCAGCTGGCTTGGAAATGAGGCTTTATTACTGGGAACCTGAGAGCTAAGAGGCTCC 121  
DB 73 TGTGGCAGTAGCTTGGGATGAGGCTCTTTCTTGTGGGAACCTGAGAGCTGAGAGGCTCC 132

QY 122 TGTGAG---CTTGTCTTAAAGTCTTAGCACTTGTGGTGGCTTGGGCTTTCACACACTGTCA 178  
DB 133 TGTGAGGCTCTGCTCTTAAACTTGTGGCAGCTTGGCGTGGCTTGGGCTTTCACACACTGTCA 192

QY 179 GACAGCTTCTGGTGGGCTTCCAGGCTTCCAGGCTTCCAGGCTTTCAGGCTTTCAGGCTTTCAGCAAGG 238  
DB 193 GACAGCTTCTGGTGGGCTTCCGSC-----TCAGGTTTGAAGCTTGGCTTTCAGCAAGG 246

QY 239 ACAGGCTCACATGAGGGGACGCCACTGATCATCTTCTGGCAGCTTCCCTTCTCTCTGCT 298  
DB 247 ACAGGCTCACATGAGGGGACGCCACTGATCATCTTCTGGCAGCTTCCCTTCTCTCTGCT 306

QY 299 TCTCTCAATGCTGTGCTGCCAGCTGTGCCGACACCCCTGTACCTGTCTTCTTGGACACCA 358  
DB 307 TCTCTCAATGCTGTATTCCAGCTGTGCCAGACCCCTGTGCTGTCTTCTTGGACACCA 366

QY 359 CCAGTGGCCACAGGGGTACCCCTGGTGTGCTGGTGTGCTGTGCTGTGCTGTGCTGTGCT 418  
DB 367 CCAGTGGCCACAGGGGTACCCCTGGTGTGCTGGTGTGCTGTGCTGTGCTGTGCTGTGCT 426

QY 419 ACGGAGGCTGGGGAGTCTCGGACCACTGCTGACGCTGCTGACGCTGCTGACGCTGCTGCT 478  
DB 427 ACGGAGGCTGGGGAGTCTCGGACCACTGCTGACGCTGCTGACGCTGCTGACGCTGCTGCT 486



```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/010,408
FILING DATE: 07-Dec-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/044,273
FILING DATE: March 19, 1998
APPLICATION NUMBER: <Unknown>
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Amy E. Mandragouras
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: MBI-004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 753 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..750
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-010-408-3

```

| Query Match           | 71.9%;          | Score 635;   | DB 15;    | Length 753; |
|-----------------------|-----------------|--|-----------|-------------|
| Best Local Similarity | 100.0%;         | Pred. No. 6.6e-176;  |           |             |
| Matches 635;          | Conservative 0; | Mismatches 0;  | Indels 0; | Gaps 0;     |
| QY                    | 249             | ATGAGGGCAGCCACCATGATCCATCTTCGGCCACTCTCTCTCTGCGCTTCTCTCAATG     | 308       |             |
| DB                    | 1               | ATGAGGGCAGCCACCATGATCCATCTTCGGCCACTCTCTCTCTGCGCTTCTCTCAATG     | 60        |             |
| QY                    | 309             | GTGTGTGCCACAGCTGTGCCGGACACCCCTGTACCTGTCTTGGACACACCCACAGTGCCCA  | 368       |             |
| DB                    | 61              | GTGTGTGCCACAGCTGTGCCGGACACCCCTGTACCTGTCTTGGACACACCCACAGTGCCCA  | 120       |             |
| QY                    | 369             | CAGGGGTACCCCTGGTGCTGGATGGCTGTGGCTGTCTGTAAGTGTGTGACGAGGCTG      | 428       |             |
| DB                    | 121             | CAGGGGTACCCCTGGTGCTGGATGGCTGTGGCTGTCTGTAAGTGTGTGACGAGGCTG      | 180       |             |
| QY                    | 429             | GGGGAGTCTTCGCACCACTCATGTCTGCGACCCACCCAGCGGCTGTGTTGTTCAGCCT     | 488       |             |
| DB                    | 181             | GGGGAGTCTTCGCACCACTCATGTCTGCGACCCACCCAGCGGCTGTGTTGTTCAGCCT     | 240       |             |
| QY                    | 489             | GGGGCAGGCCCTGGCGGCCATGGGGCTGTGTGTCTCTTGATGAGGATGACGCTAGCTGT    | 548       |             |
| DB                    | 241             | GGGGCAGGCCCTGGCGGCCATGGGGCTGTGTGTCTCTTGATGAGGATGACGCTAGCTGT    | 300       |             |
| QY                    | 549             | GAGGTGAATGGCCGACAGTACCTGGATGGAGAGACCTTTAAACCCCAATTCGACGGTCTCTG | 608       |             |
| DB                    | 301             | GAGGTGAATGGCCGACAGTACCTGGATGGAGAGACCTTTAAACCCCAATTCGACGGTCTCTG | 360       |             |
| QY                    | 609             | TGCCCTCTGATGACGGTGGGCTTCACTCTGCTGCGCTGTGCACTGAGGATGTGGGGCTG    | 668       |             |
| DB                    | 361             | TGCCCTCTGATGACGGTGGGCTTCACTCTGCTGCGCTGTGCACTGAGGATGTGGGGCTG    | 420       |             |
| QY                    | 669             | CCCACTGGGACTGCCACCGCCCAAGAGAAATCAGGTGCCAGGAAAGTGTGCCCCGAG      | 728       |             |
| DB                    | 421             | CCCACTGGGACTGCCACCGCCCAAGAGAAATCAGGTGCCAGGAAAGTGTGCCCCGAG      | 480       |             |
| QY                    | 729             | TGGGTATGTACAGGGAGGTACACCGCGCATCCAGGCTCCACGCGCAAGACACCA         | 788       |             |
| DB                    | 481             | TGGGTATGTACAGGGAGGTACACCGCGCATCCAGGCTCCACGCGCAAGACACCA         | 540       |             |
| QY                    | 789             | CTTCTGCCCTTGTCACCTCTGCTCTGTGCTGCTCTTGTCCAAATTTGGAGCACAGCC      | 848       |             |

```

Db      541 CTTTCTCCCTTGTCACCTGCGCTGCTGATCCTTGTGCATAATTGGAGCACGCC   600
Qy      849 TGGGGCCCCGTCAACCACCTGTGGCTGGGCAT   883
Db      601 TGGGGCCCCGTCAACCACCTGTGGCTGGGCAT   635

RESULT 5
US-10-010-408-12
; Sequence 12, Application US/10010408
; Publication No. US20020165185A1
; GENERAL INFORMATION:
; APPLICANT: John J. Castellet, Jr.
; TITLE OF INVENTION: NO. US20020165185A1el Heparin-Induced CCN-Like Molecules
; and Uses Therefor
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/010,408
; FILING DATE: 07-Dec-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/044,273
; FILING DATE: March 19, 1998
; APPLICATION NUMBER: <Unknown>
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Amy E. Mandragoras
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: MBI-004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 681 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..681
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
--JS-10-010-408-12

```

[illegible]

|    |     |   |     |
|----|-----|---|-----|
| Qy | 498 | CCTGGCGGCATGGGGCTGTGTGTCCTTTGGATGAGGATGACGGTAGCTGTGAGSTGAAT     | 557 |
| Db | 181 |   |     |
| Db | 181 | CCTGGCGGCATGGGGCTGTGTGTCCTTTGGATGAGGATGACGGTAGCTGTGAGSTGAAT     | 240 |
| Qy | 558 | GCGCCACGGTACCTGGATGGAGAGACACTTTAAACCCCAATTGCAGGGTCTCTGTGCCGCTGT | 617 |
| Db | 241 |   |     |
| Db | 241 | GCGCCACGGTACCTGGATGGAGAGACACTTTAAACCCCAATTGCAGGGTCTCTGTGCCGCTGT | 300 |
| Qy | 618 | GATGACGGTGGCTTCACTGCTGCCGCTGTGTGACGTGAGGATGTGGGGTGCCACAGCTGG    | 677 |
| Db | 301 |   |     |
| Db | 301 | GATGACGGTGGCTTCACTGCTGCCGCTGTGTGACGTGAGGATGTGGGGTGCCACAGCTGG    | 360 |
| Qy | 678 | GACTGCCACGCCCCACAGAGAATAAGGTGCCAGGAAAGTGTGCCCCGAGTGGGTATGT      | 737 |
| Db | 361 |   |     |
| Db | 361 | GACTGCCACGCCCCACAGAGAATAAGGTGCCAGGAAAGTGTGCCCCGAGTGGGTATGT      | 420 |
| Qy | 738 | GACCAGGGAGTGACACCGGCGCATCAGCGTCCACGGCGCAAGGACACCAACTTTCGTGCC    | 797 |
| Db | 421 |   |     |
| Db | 421 | GACCAGGGAGTGACACCGGCGCATCAGCGTCCACGGCGCAAGGACACCAACTTTCGTGCC    | 480 |
| Qy | 798 | CTTTGTCACTCTCGCTCTGTGATGCTCTTGTCCAAAATTGGAGACACAGCCTGGGGCCCC    | 857 |
| Db | 481 |   |     |
| Db | 481 | CTTTGTCACTCTCGCTCTGTGATGCTCTTGTCCAAAATTGGAGACACAGCCTGGGGCCCC    | 540 |
| Qy | 858 | TGCTCAACACACTGTGGGCTGGGCAT                                      | 883 |
| Db | 541 |   |     |
| Db | 541 | TGCTCAACACACTGTGGGCTGGGCAT                                      | 566 |

## RESULT 6

```

US-09-915-582-30
; Sequence 30, Application US/03915582
; Patent No. US20020120103A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 17 Human Secreted Pro
; FILE REFERENCE: P5723P1
; CURRENT APPLICATION NUMBER: US/09/915,582
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: PCT/US01/01431
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/231,968
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
; LENGTH: 1337
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1337)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-915-582-30

```

|    | Query Match           | 49.8%;   | Score 440;         | DB 11; | Length 1337; |
|----|-----------------------|--|--------------------|--------|--------------|
|    | Best Local Similarity | 78.3%;   | Pred. No. 8e-119;  |        |              |
|    | Matches 539;          | Conservative   | 3; Mismatches 138; | Indels | 8; Gaps 1    |
| QY | 196                   | CTCCAGGCCCTCACCTTCAGGTTTGAAGCTGGGTCCACAGGACACGGTGACATGAGG    | 255                |        |              |
|    |                       |  |                    |        |              |
| Db | 2                     | CTTCACAGTTTCACCTTCAGGCTCAAAACTGGSTGTCA                       |                    |        |              |
|    |                       |  |                    |        |              |
| QY | 256                   | GCAGCCCACATGATTCATCTCTGGCCACTTCCTTCCTTCGCTTCTCTCAATGGTGTGG   | 315                |        |              |
|    |                       |  |                    |        |              |
| Db | 54                    | GCACACGNAAGACCACCTCTCGGCTTCTCCCCCTCCTTCGCTCCTCTCAAAGGTCGGTA  | 113                |        |              |
|    |                       |  |                    |        |              |
| QY | 316                   | CCGAGCTGTGCGCGGACACCCCTGTACCTGTCTCGGACACACCCCGAGTGCCACAGGGGG | 375                |        |              |
|    |                       |  |                    |        |              |

|     |   |     |
|-----|---|-----|
| 114 | CCCAGCTGTGCCCGACACCATGTACCTGCCCCCTGGCCACCTCCCGCATGCCGCTGGGAG  | 173 |
| 376 | TACCCCTGGTGCTGGATGGCTGTGGCTGCTCTAAAGTGTGTGCACGGAGGTGGGGGAGT   | 435 |
| 174 | TACCCCTGGTGCTGGATGGCTGTGGCTGCTCCCGGGTATGTGCACGGCGGTGGGGGAGC   | 233 |
| 436 | CTTGCGACCACTTCGATGTCGACACCCACCCAGCGGCCCTGGTTGTGACGCTGGGGCAG   | 495 |
| 234 | CTTGCGACCAACTCCAGTCTGCGACGCGCACGCCAGGGCCTGGTCTGCCAGCCCGGGGCG  | 293 |
| 496 | GCCCTGGCGGCATGGGGCTGTGTGTCTTCTTGGATGAGGATGACGGTAGCTGTGAGGTGA  | 555 |
| 294 | GACCCGTGMCGGGGGGCCCTGTGCTCTTTGGCAGAGGACGACAGCACTGTGAGGTGA     | 353 |
| 556 | ATGCCCGCAGGTACTCGATGGAGAGACCTTTTAAACCAATTGCAGGGTCTCTGTGCCGCT  | 615 |
| 354 | ACGCCGCCCTGTATCGGGAAGGGGAGACCTTCCAGCCCACTGCAGCACTCCGCTGCCGCT  | 413 |
| 616 | GTGATGACGGTGCGCTTACCTGCGCTCCGCGCTGCAGTGAGATGTGCGGCTGCCCAAGCT  | 675 |
| 414 | GCGAGGACGGGGGCTTACCTTGCCTGCGCGCTGTGCAGCGAGATGTGCGGCTGCCCAAGCT | 473 |
| 676 | GGACTGCCCAACGCCCAAGAGAATACAGGTGCCAGGAAAGTGTCTGCCCGAGTGGGTAT   | 735 |
| 474 | GGACTGCCCAACCCACGAGGGTCTGAGGTCCTGGGCAAGTGTCTGAGTGGGTCT        | 533 |
| 736 | GTGACCAAGGAGTGACACCGCGATCCAGCGCTCCACGGCGCAAGGACACCAACTTTCTTG  | 795 |
| 534 | GCGGCCAAGGAGGGGGACTGGGGACCCAGCCCTTCCAGCCCAAGAGACCCCAAGTTTCTG  | 593 |
| 796 | CCCTTGTACCTCCCTCTCTCTGATGCTCTTGTCTCCAAATTTGGAGCACAGCTGGGGC    | 855 |
| 594 | GCCTTGTCTTCCCTGCCCCCTGGTGTCCCTGCCCAAGATGGAGCACGGCCTGGGGAC     | 653 |
| 856 | CTTGCTCAACCACTGTGGGCTGGGCAT                                   | 883 |
| 654 | CTCTCTCGAACCACTGTGGGCTGGGCAT                                  | 681 |

## RESULT 7

```

RES001 7
US-09-915-582-14
; Sequence 14, Application US/09915582
; Patent No. US20020120103A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 17 Human Secreted Proteins
; FILE REFERENCE: PST23P1
; CURRENT APPLICATION NUMBER: US/09/915,582
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: PCT/US01/01431
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/231,968
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 1352
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-915-582-14

```

```

Query Match      49.98;   Score 439.8;   DB 11;   Length 1352;
Best Local Similarity 78.4%;   Pred. No. 9.1e-119;
Matches 543;   Conservative 0;   Mismatches 142;   Indels 8;   Gaps 1;

Qy 191 GTGGCTCCACGGCCTCACGCTTCAGGTTTGAAGCTGGCTCCCAAGGGACACGGTGACAT 250
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3 GTCCGTTTCAGATTTCACCTTCAGGCTCAAGGCTGGCTCTGCA-----GGGGACAT 54

```





[illegible]

QY 843 ACAGCCTGGGGCCCTGTCTCAACACCTGTGGGTGGGCAT 883  
DB 604 ACGGCTGGGGACCGTCTGCACACCTGTGGGTGGGCAT 644

## RESULT 9

US-10-146-726-319  
; Sequence 319, Application US/10146726  
; Publication No. US20030129690A1

## GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Collin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; TITLE OF INVENTION: ACIDS ENCODING THE SAME  
; FILE REFERENCE: P3330R1C308  
; CURRENT APPLICATION NUMBER: US/10/146,726  
; CURRENT FILING DATE: 2002-05-15  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 319  
; LENGTH: 1266  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-146-726-319

Query Match 49.2%; Score 434.6; DB 14; Length 1266;

Best Local Similarity 79.9%; Pred. No. 3e-117; Indels 0; Gaps 0;  
Matches 512; Conservative 0; Mismatches 129;

QY 243 GGTGACATGAGGGGAGCCCACTGATCCATCTTGTGGCCACTTCTCTCTGCTTCTC 302  
DB 4 GGGGACATGAGAGGCACACCGACCCACCTCTCTGCTTCTCTCTCTCTCTCTC 63  
QY 303 TCAATGTTGTGCGGACACCCCTGTACCTGTCTCTGACACCCACCCAG 362  
DB 64 TCAAGGTGCGTACCGAGCTGTGCCGACACCATGTACCTGCCCTGCCACCTCCCGA 123  
QY 363 TGCCACAGGGGTGACCCCTGGTGGATGGCTGTCTGTAAGTGTGTGCACGG 422  
DB 124 TGCCCGTGGGAGTACCCCTGGTGGATGGCTGTCTGCTGCTGGGTATGTGCACGG 183  
QY 423 AGGCTGGGGAGTCTCGGACACCTGTCTGCGACCCAGCCAGCGGCTGTGTTGT 482  
DB 184 CGGCTGGGGAGCCCTGGGACCACTCCACCTCTGCGACGCGCAGGCGCTGGTCTGC 243  
QY 483 CAGCCTGGGGAGCCCTGGCGGCATGGGCTGTGTCTCTTGGATGAGGATGACGGT 542  
DB 244 CAGCCCGGGGAGGACCGGTGGCGGGGCCCTGTGCCCTCTTGGGAGAGGACAGCAGC 303  
QY 543 AGCTGTGAGGTGAATGTCGCGCAGGTACCTGGATGGAGAGACCTTTAAACCAATGACGG 602  
DB 304 AGCTGTGAGGTGAACGCGGCTCTATGTCGAAGGGGAGAGCTTCCAGCCCACTGACAGC 363  
QY 603 GTCCTGTCCGCTGTGATGACGGTGGCTTACCTGCCCTGCCCTGTGTCAGTGAGGATGTG 662  
DB 364 ATCCGCTGCCCTGGGAGGAGCGGCTTACCTTGCCTGCGCTGTGTCGACGAGGATGTG 423

QY 663 CGGCTGCCCAAGTGGGACTGCCACGCCCAAGAGATAACAGGTGCCAGGAAAGTGTGC 722  
DB 424 CGGCTGCCCAAGTGGGACTGCCACGCCCAAGAGGTGCGAGGTCTCTGGCAAGTGTGC 483  
QY 723 CCCGAGTGGGTATGTGACCAAGGAGTGACACGGGCGATCCAGCGCTCCACGGCGCAAGA 782  
DB 484 CCTGAGTGGGTGTCGGGCCAAGGAGGAGTGGGGACCCAGCCCTTCCAGCCCAAGA 543  
QY 783 CACCAACTTCTGCCCCTTGTTCACCTCTCTGATGCTCTCTTGTCCAAATTTGGAGC 842  
DB 544 CCCAGTCTTCTGCGCTTGTCTCTCTCTGCTGCTCTCTCTCTCTCTCTCTCTCTCT 603  
QY 843 ACAGCCTGGGGCCCTGCTCTCAACACCTGTGGGTGGGCAT 883  
DB 604 ACGGCTGGGACCTGCTCTGACACCTGTGGGTGGGCAT 644

## RESULT 10

US-10-146-727-319  
; Sequence 319, Application US/10146727  
; Publication No. US20030129691A1

## GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Collin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; TITLE OF INVENTION: ACIDS ENCODING THE SAME  
; FILE REFERENCE: P3330R1C312  
; CURRENT APPLICATION NUMBER: US/10/146,727  
; CURRENT FILING DATE: 2002-05-15  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 319  
; LENGTH: 1266  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-146-727-319

Query Match 49.2%; Score 434.6; DB 14; Length 1266;

Best Local Similarity 79.9%; Pred. No. 3e-117; Indels 0; Gaps 0;  
Matches 512; Conservative 0; Mismatches 129;

QY 243 GGTGACATGAGGGGAGCCCACTGATCCATCTTGTGGCCACTTCTCTCTGCTTCTC 302  
DB 4 GGGGACATGAGAGGCACACCGACCCACCTCTCTGCTTCTCTCTCTCTCTCTC 63  
QY 303 TCAATGTTGTGCGGACACCCCTGTACCTGTCTCTTGGACACCCACCCAG 362  
DB 64 TCAAGGTGCGTACCGAGCTGTGCCGACACCATGTACCTGCCCTGCCACCTCCCGA 123  
QY 363 TGCCACAGGGGTGACCCCTGGTGGATGGCTGTCTGTAAGTGTGTGCACGG 422  
DB 124 TGCCCGTGGGAGTACCCCTGGTGGATGGCTGTCTGCTGCTGGGTATGTGCACGG 183  
QY 423 AGGCTGGGGAGTCTCGGACACCTGTCTGCGACCCAGCCAGCGGCTGTGTTGT 482  
DB 184 CGGCTGGGGAGCCCTGGGACCACTCCACCTCTGCGACGCGCAGGCGCTGGTCTGC 243  
QY 483 CAGCCTGGGGAGGCGCTTGGCGGCATGGGGCTGTGTCTCTTGGATGAGGATGACGGT 542

```
Db 244 CAGCCCGGGGAGGACCCGGTGGCCGGGGGCCCTGTGCTCTTTGGCAGAGGACGACAGC 303
Qy 543 AGCTGTGAGTGAATGGCCGAGGTACTGTGATGGAGAGACCTTTAAACCCCAATTGCAGG 602
Db 304 AGCTGTGAGTGAACGGCCGCTGTATCGGGAAGGGAGACCTTCCAGCCCACTGCAGC 363
Qy 603 GTCTGTGCGCTGTGATGAGCGTGTACCTGCTGCGCTGTGCTGCTGCTGCTGCTGCTGCTG 662
Db 364 ATCCGCTGCGCTGCGAGGACGGGGCTTACCTGCTGCGCTGTGCTGCTGCTGCTGCTGCTG 423
Qy 663 CGGCTGCGGCTGCGGACTGCCACGCGCCCAAGAGATACAGTGCAGGAAATGCTGTC 722
Db 424 CGGCTGCGGCTGCGGACTGCCACGCGCCCAAGAGATACAGTGCAGGAAATGCTGTC 483
Qy 723 CCCGAGTGGGTATGTGACAGGAGTGCACACCGCGGATCCAGCGCTCCACGGCCCAAGGA 782
Db 484 CCGTGTGAGTGTGCGGCGCAGGAGGAGGACTGGGGACCCACGCCCTTCCAGCCCAAGGA 543
Qy 783 CACCAACTTTTGTGCGCTGTGATGAGCGTGTACCTGCTGCTGCTGCTGCTGCTGCTGCTG 842
Db 544 CCCCAGTTTTCTGCGCTTGTCTCTTCCCTGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTG 603
Qy 843 ACAGCTGGGGCCCTGCTCAACACCTGTGGGCTGGGCAT 883
Db 604 ACGGCTGGGGACCCCTGCTCGACACCTGTGGGCTGGGCAT 644
```

## RESULT 11

US-10-146-788-319

```
; Sequence 319, Application US/10146788
; Publication No. US20030129693A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tamas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C322
; CURRENT APPLICATION NUMBER: US/10/146,788
; CURRENT FILING DATE: 2002-05-15
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 319
; LENGTH: 1266
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-146-788-319
```

```
Query Match 49.2%; Score 434.6; DB 14; Length 1266;
Best Local Similarity 79.9%; Pred. No. 3e-117;
Matches 512; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

Qy 243 GGTGACATGAGGGGACCCAGTCCATCTTCTGCGGCACTTCTCTTCTGCTGCTGCTGCTGCTG 302
Db 4 GGGGACATGAGGACACCGAGACCCACCTCTCTGCGCTTCTCTCTCTCTCTCTCTCTCTCTCTC 63
Qy 303 TCAATGTGTGTCGCGGAGTGTGCGGACACCCCTGTACCTGTCTCTGAGACACCAACCCAG 362
```

```
Db 64 TCAAAGTGGGTACCCAGCTGTGCCCGACACCATGTACCTGCCCCCTGGCCACCTCCCCGGA 123
Qy 363 TGCCACACAGGGGTACCCCTGGTGTGATGGTGTGGCTGTCTGCTGCTGCTGCTGCTGCTGCTG 422
Db 124 TGCCCGCTGGGAGTACCCCTGGTGTGATGGTGTGGCTGTGGCTGTGCTGCTGCTGCTGCTG 183
Qy 423 AGCTGTGGGGAGTCTCGGACCACTGTCTGCGACCCCAAGAGATACAGTGCAGGAAATGCTG 482
Db 184 CGGCTGGGGAGCCCTGCGACCAACTCCACGTCTGCGACGCGCAGCCAGGGCTGTGCTGTC 243
Qy 483 CAGCCTGGGCGAGCCCTGCGGCGCATGCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 542
Db 244 CAGCCCGGGGACAGCCCGTGGCCGGGGGCCCTGTGCTCTTTGGCAGAGGACGACAGC 303
Qy 543 AGCTGTGAGTGAATGGCCGAGGTACTGTGATGGAGAGACCTTTAAACCCCAATTGCAGG 602
Db 304 AGCTGTGAGTGAACGGCCGCTGTATCGGGAAGGGAGACCTTCCAGCCCACTGCAGC 363
Qy 603 GTCTGTGCGCTGTGATGAGCGTGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 662
Db 364 ATCCGCTGCGCTGCGAGGACGGGGCTTACCTGCTGCGCTGTGCTGCTGCTGCTGCTGCTG 423
Qy 663 CGGCTGCGGCTGCGGACTGCCACGCGCCCAAGAGATACAGTGCAGGAAATGCTGTC 722
Db 424 CGGCTGCGGCTGCGGACTGCCACGCGCCCAAGAGATACAGTGCAGGAAATGCTGTC 483
Qy 723 CCCGAGTGGGTATGTGACAGGAGTGCACACCGCGGATCCAGCGCTCCACGGCCCAAGGA 782
Db 484 CCGTGTGAGTGTGCGGCGCAGGAGGAGGACTGGGGACCCACGCCCTTCCAGCCCAAGGA 543
Qy 783 CACCAACTTTTGTGCGCTGTGATGAGCGTGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 842
Db 544 CCCCAGTTTTCTGCGCTTGTCTCTTCCCTGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTG 603
Qy 843 ACAGCTGGGGCCCTGCTCAACACCTGTGGGCTGGGCAT 883
Db 604 ACGGCTGGGGACCCCTGCTCGACACCTGTGGGCTGGGCAT 644
```

## RESULT 12

```
US-10-152-380-319
; Sequence 319, Application US/10152380
; Publication No. US20030129694A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tamas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C397
; CURRENT APPLICATION NUMBER: US/10/152,380
; CURRENT FILING DATE: 2002-05-21
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 319
; LENGTH: 1266
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-152-380-319
```

[illegible]

RESULT 14  
US-10-028-072-319  
; Sequence 319, Application US/10028072  
; Publication No. US2003000431A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: Desnoyers, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.

APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang  
TITLE OF INVENTION:  
FILE REFERENCE:  
CURRENT APPLICATION NUMBER: US/10/028,072  
CURRENT FILING DATE: 2001-12-19  
PRIOR APPLICATION NUMBER: 60/045911  
PRIOR FILING DATE: 1997-06-18  
PRIOR APPLICATION NUMBER: 60/056974  
PRIOR FILING DATE: 1997-08-26  
PRIOR APPLICATION NUMBER: 60/059113  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/059115  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/059117  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/059122  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/059184  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/059263  
PRIOR FILING DATE: 1997-09-18  
PRIOR APPLICATION NUMBER: 60/059352  
PRIOR FILING DATE: 1997-09-19  
PRIOR APPLICATION NUMBER: 60/059588  
PRIOR FILING DATE: 1997-09-19  
PRIOR APPLICATION NUMBER: 60/059836  
PRIOR FILING DATE: 1997-09-24  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/062285  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/062287  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/062814  
PRIOR FILING DATE: 1997-10-24  
PRIOR APPLICATION NUMBER: 60/062816  
PRIOR FILING DATE: 1997-10-24  
PRIOR APPLICATION NUMBER: 60/063045  
PRIOR FILING DATE: 1997-10-24  
PRIOR APPLICATION NUMBER: 60/063082  
PRIOR FILING DATE: 1997-10-31  
PRIOR APPLICATION NUMBER: 60/063127  
PRIOR FILING DATE: 1997-10-24  
PRIOR APPLICATION NUMBER: 60/063327  
PRIOR FILING DATE: 1997-10-27  
PRIOR APPLICATION NUMBER: 60/063329  
PRIOR FILING DATE: 1997-10-27  
PRIOR APPLICATION NUMBER: 60/063550  
PRIOR FILING DATE: 1997-10-28  
PRIOR APPLICATION NUMBER: 60/063561  
PRIOR FILING DATE: 1997-10-28  
PRIOR APPLICATION NUMBER: 60/063704  
PRIOR FILING DATE: 1997-10-29  
PRIOR APPLICATION NUMBER: 60/063733  
PRIOR FILING DATE: 1997-10-29  
PRIOR APPLICATION NUMBER: 60/063735  
PRIOR FILING DATE: 1997-10-29  
PRIOR APPLICATION NUMBER: 60/063738  
PRIOR FILING DATE: 1997-10-29  
PRIOR APPLICATION NUMBER: 60/063755  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/064248  
PRIOR FILING DATE: 1997-11-03  
PRIOR APPLICATION NUMBER: 60/064809  
PRIOR FILING DATE: 1997-11-07  
PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065846  
PRIOR FILING DATE: 1997-11-17  
PRIOR APPLICATION NUMBER: 60/066364  
PRIOR FILING DATE: 1997-11-21  
PRIOR APPLICATION NUMBER: 60/066453  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/066511  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/066770  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/069212  
PRIOR FILING DATE: 1997-12-11  
PRIOR APPLICATION NUMBER: 60/069278  
PRIOR FILING DATE: 1997-12-11  
PRIOR APPLICATION NUMBER: 60/069334  
PRIOR FILING DATE: 1997-12-11  
PRIOR APPLICATION NUMBER: 60/069694  
PRIOR FILING DATE: 1997-12-16  
PRIOR APPLICATION NUMBER: 60/072320  
PRIOR FILING DATE: 1998-01-23  
PRIOR APPLICATION NUMBER: 60/073612  
PRIOR FILING DATE: 1998-02-04  
PRIOR APPLICATION NUMBER: 60/074086  
PRIOR FILING DATE: 1998-02-09  
PRIOR APPLICATION NUMBER: 60/074092  
PRIOR FILING DATE: 1998-02-09  
PRIOR APPLICATION NUMBER: 60/077791  
PRIOR FILING DATE: 1998-03-12  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/079294  
PRIOR FILING DATE: 1998-03-25  
PRIOR APPLICATION NUMBER: 60/079663  
PRIOR FILING DATE: 1998-02-27  
PRIOR APPLICATION NUMBER: 60/079728  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/080165  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/081203  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081229  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081695  
PRIOR FILING DATE: 1998-04-14  
PRIOR APPLICATION NUMBER: 60/081817  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081818  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/082999  
PRIOR FILING DATE: 1998-04-24  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/083545  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084627  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084637  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/085149  
PRIOR FILING DATE: 1998-05-12  
PRIOR APPLICATION NUMBER: 60/085323  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085338  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085339  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085579  
PRIOR FILING DATE: 1998-05-15

|   |                                     |
|---|-------------------------------------|
| ; | PRIOR APPLICATION NUMBER: 60/085697 |
| ; | PRIOR FILING DATE: 1998-05-15       |
| ; | PRIOR APPLICATION NUMBER: 60/085704 |
| ; | PRIOR FILING DATE: 1998-05-15       |
| ; | PRIOR APPLICATION NUMBER: 60/086414 |
| ; | PRIOR FILING DATE: 1998-05-22       |
| ; | PRIOR APPLICATION NUMBER: 60/086430 |
| ; | PRIOR FILING DATE: 1998-05-22       |
| ; | PRIOR APPLICATION NUMBER: 60/087106 |
| ; | PRIOR FILING DATE: 1998-05-28       |
| ; | PRIOR APPLICATION NUMBER: 60/088026 |
| ; | PRIOR FILING DATE: 1998-06-04       |
| ; | PRIOR APPLICATION NUMBER: 60/088730 |
| ; | PRIOR FILING DATE: 1998-06-10       |
| ; | PRIOR APPLICATION NUMBER: 60/088741 |
| ; | PRIOR FILING DATE: 1998-06-10       |
| ; | PRIOR APPLICATION NUMBER: 60/088810 |
| ; | PRIOR FILING DATE: 1998-06-10       |
| ; | PRIOR APPLICATION NUMBER: 60/088858 |
| ; | PRIOR FILING DATE: 19/98-06-11      |
| ; | PRIOR APPLICATION NUMBER: 60/089532 |
| ; | PRIOR FILING DATE: 1998-06-17       |
| ; | PRIOR APPLICATION NUMBER: 60/089599 |
| ; | PRIOR FILING DATE: 1998-06-17       |
| ; | PRIOR APPLICATION NUMBER: 60/089907 |
| ; | PRIOR FILING DATE: 1998-06-18       |
| ; | PRIOR APPLICATION NUMBER: 60/089947 |
| ; | PRIOR FILING DATE: 1998-06-19       |
| ; | PRIOR APPLICATION NUMBER: 60/090349 |
| ; | PRIOR FILING DATE: 1998-06-23       |
| ; | PRIOR APPLICATION NUMBER: 60/090429 |
| ; | PRIOR FILING DATE: 1998-06-24       |
| ; | PRIOR APPLICATION NUMBER: 60/090445 |
| ; | PRIOR FILING DATE: 1998-06-24       |
| ; | PRIOR APPLICATION NUMBER: 60/090538 |
| ; | PRIOR FILING DATE: 1998-06-24       |
| ; | PRIOR APPLICATION NUMBER: 60/090863 |
| ; | PRIOR FILING DATE: 1998-06-26       |
| ; | PRIOR APPLICATION NUMBER: 60/091360 |
| ; | PRIOR FILING DATE: 1998-07-01       |
| ; | PRIOR APPLICATION NUMBER: 60/091519 |
| ; | PRIOR FILING DATE: 1998-07-02       |
| ; | PRIOR APPLICATION NUMBER: 60/091982 |
| ; | PRIOR FILING DATE: 1998-07-07       |

[illegible]

|    |     |   |     |
|----|-----|---|-----|
| Qy | 603 | GTCTGTGCCCGCTGTGTATGACGGTGGCTTACCTACCTGCCCTGTGCAGTGCAGATGTG         | 662 |
|    |     |   |     |
| Db | 364 | ATCCGTGCCCGCTGGAGGACGGGGCTTCACTGCGTCCCGCTGTGCAGCGAGGATGTG           | 423 |
|    |     |   |     |
| Qy | 663 | CGGCTGCCACGCTGGGACTGCCACGCCCCCAAGAGAAATACAGGTGCCAGGAAAGTCTGC        | 722 |
|    |     |   |     |
| Db | 424 | CGGCTGCCACGCTGGGACTGCCACCCAGAGGGTCTGGGCAAGTCTGTGC                   | 483 |
|    |     |   |     |
| Qy | 723 | CCCAGTGGGTATGTGACAGGGAGTGCACACGGCGATCCACGCGCTCCACGCGCAAGGA          | 782 |
|    |     |   |     |
| Db | 484 | CCTCAGTGGGTGTGCGGCCAAGAGGGGACTGGGGACCCAGCCCCCTTCCAGCCCAAGGA         | 543 |
|    |     |   |     |
| Qy | 783 | CACCAACTTCTGCCCTGTGCATCTCCTGCCCTGTGCTGTGATGCTCTGTGTCCTCAAAATTTGGAGC | 842 |
|    |     |   |     |
| Db | 544 | CCCCAGTTTCTGGCGCTGTCTCTTCCCTGGCCCCCTGTGCTGCCCAAGATGGAGC             | 603 |
|    |     |   |     |
| Qy | 843 | ACAGCTGGGGCCCTGCTCAACACACCTGTGGCTGGGCAT                             | 883 |
|    |     |   |     |
| Db | 604 | ACGGCTGGGGACCCCTGCTGACCACCTGTGGCTGGGCAT                             | 644 |
|    |     |   |     |

RESULT 15

US-10-121-049-319

; Sequence 319, Application US/10121049

; Publication No. US20030022239A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Beresini, Maureen

; APPLICANT: DeForge, Laura

; APPLICANT: Desnoyers, Luc

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Sherwood, Steven

; APPLICANT: Smith, Victoria

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Watanabe, Colin K

; APPLICANT: Wood, William

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE OF INVENTION: ACTOS ENCODING THE SAME

; FILE REFERENCE: P3330R1G17

; CURRENT APPLICATION NUMBER: US/10121,049

; CURRENT FILING DATE: 2002-04-12

Prior Application removed - See File Wrapper or Palm

; NUMBER OF SEQ ID NOS: 550

; SEQ ID NO 319

; LENGTH: 1266

; TYPE: DNA

; ORGANISM: Homo Sapien

; SS-10-121-049-319

|    | Query Match           | 49.2%   | Score 434.6      | DB 15    | Length 1266 |
|----|-----------------------|---|------------------|----------|-------------|
|    | Best Local Similarity | 79.9%   | Pred. No. 3e-117 |          |             |
|    | Matches 512           | Conservative 0  | Mismatches 129   | Indels 0 | Gaps 0      |
| Qy | 243                   | GGTGACATGAGGGCAGCCCACTGATCCATCTTCTGGCCACTTCTTCCCTCTGCCTTCTC   | 302              |          |             |
| Db | 4                     | GGGACATGAGGACACCGAGACCACCTCTCGGCTTCTCCCTTCTTGGCTCTCTC         | 63               |          |             |
| Qy | 303                   | TCAATGGTGTGTGCCAGCTGTGCGGACACCCCTGTACCTGTCTTGGACACCACCCCAG    | 362              |          |             |
| Db | 64                    | TCAAAGTGTGCTACCCAGCTGTGCCGACACCATGTACCTGCCCTTGCCACCTCCCCGA    | 123              |          |             |
| Qy | 363                   | TGCCACAGGGGTACCCCTGTGTGCTGGATGGCTGTGGCTGTCTAAAGTGTGTGCACGG    | 422              |          |             |
| Db | 124                   | TGCCCGGTGGAGTACCCCTGTGTGCTGGATGGCTGTGGCTGTCTGCCGGGTATGTGCACGG | 183              |          |             |
| Qy | 423                   | AGGCTGGGGAGTCTCTGCGACCACTGCATGTCTTGGCACCCCAAGGGCCTCGTGTGT     | 482              |          |             |

|     |    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |     |
|-----|----|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|-----|
| 184 | Db |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | CGGCTGGGGAGCCCTCGACCAACTCCACGCTCTGCAGCCAGCCAGGGCTGGTCTGTC      | 243 |
| 483 | Qy |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | CAGCTGGGCGAGCCCTGCGGGCCATGGGGCTGTGTGCTCTCTTGGATGAGGATGACGGT    | 542 |
| 244 | Db |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | CAGCCGGGCGAGACCCTGTGGCCGGGGGCCCTGTGCCCTCTTTGGCAGAGGACGACAGC    | 303 |
| 543 | Qy |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | AGCTGTGAGTGAATGSCCGCAGGTACCTGGATGGAGAGACCTTTAAACCAATTCAGG      | 602 |
| 304 | Db |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | AGCTGTGAGTGAACGGCCCTGTATCGGGAAGGGAGACCTTCCAGCCCCACTGCAGC       | 363 |
| 603 | Qy |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | GTCTGTGCCGCTGTGATGAGCGGTGGCTTCACTTCGCTGCGCTGTGCAGTGAGGATGTG    | 662 |
| 364 | Db |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | ATCCGCTGCCGCTCGAGGAGCGCGGCTTCACTTCGCTGCGCTGTGCAAGCGAGGATGTG    | 423 |
| 663 | Qy |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | CGGTCGCCAGCTGGGACTGCCACGCCCCCAAGAGATACAGGTGCCAGGAAGTGTGTC      | 722 |
| 424 | Db |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | CGGTCGCCAGCTGGGACTGCCCCACCCAGAGGGTCTGAGGCTCTGGGGCAAGTGTGTC     | 483 |
| 723 | Qy |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | CCGAGTGGGTATGTGACCAGGGAGTGACACCGGCGATCCAGCGGTCTCCACGGCGCAAGGA  | 782 |
| 484 | Db |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | CTGTAGTGGGTGTGCGGCCAAGGAGGGGACTGGGGACCCAGCCCTTCCAGCCCAAGGA     | 543 |
| 783 | Qy |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | CACCAACTTTCTGCCCTTGTCACTTCCTGCTCTCTGTGATGCTCCTTGTCCAAATTTGGAGC | 842 |
| 544 | Db |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | CCCAAGTTTCTGSCCTTGTCTCTCTCCTGCCCTGGTGCCTGCCAGATGGAGC           | 603 |
| 843 | Qy |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | ACAGCTGGGGCCCTGCTCAACACCTGTGGGTGGGCAT                          | 883 |
| 604 | Db |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | ACGGCTTGGGAGCCCTGCTGCACCAACCTGTGGCTGGGCAT                      | 644 |

Search completed: July 28, 2003, 21:31:53  
Job time : 218.668 secs











|    | Matches | 485; | Conservative  | 0;  | Mismatches | 120; | Indels | 0; | Gaps | 0; |
|----|---------|------|---|-----|------------|------|--------|----|------|----|
| Qy | 276     | CTG  | GCACACTTCCCTTCCTCCTCAATGGTGTTGCCAGCTGTGCGGACACCC        | 335 |            |      |        |    |      |    |
| Dd | 2       | CTG  | GCCTTCTCCCTTCCTCTGCCTCTCTCAAAGTGCATCCCAAGCTGTGCCGACACCA | 61  |            |      |        |    |      |    |
| Qy | 336     | TGT  | ACCTGTCTTGGACACCACCCACAGTGCACAGGGGTACCCCTGGTCTGGATGGC   | 395 |            |      |        |    |      |    |
| Dd | 62      | TGT  | ACCTTGCCTTGGCCACCTCCCGCATGCCCGCTGGAGTACCCCTGGTCTGGATGGC | 121 |            |      |        |    |      |    |
| Qy | 396     | TGT  | GGCTGCTCTAAAGTGTGTGCACGAGGCTGGGGAGTCTCGGACCACTCTCATGTC  | 455 |            |      |        |    |      |    |
| Dd | 122     | TGT  | GGCTGCTCCCGGTATGTGCACGCGCTGGGGAGCCCTCGGACCACTCCAGTC     | 181 |            |      |        |    |      |    |
| Qy | 456     | TGG  | ACCCCAGCAGGGCCTGTTTTCAGCTGGGGCAGGCCCTGGCGGCCATTGGGCT    | 515 |            |      |        |    |      |    |
| Dd | 182     | TGG  | CACCCAGCCAGGGCCTGGTGTGCACGCCGGGCAGGACCGGTGGCCGGGGGCC    | 241 |            |      |        |    |      |    |
| Qy | 516     | GTG  | TGCTCTTTGGATGAGGTAGCGTAGCTGTGAGTGAATGCCCGCAGGTACTCGAT   | 575 |            |      |        |    |      |    |
| Dd | 242     | CTG  | TGCTCTTTGGCAGAGGACGACAGACTGTGAGTGAACGCCCGCTGTATCGGAA    | 301 |            |      |        |    |      |    |
| Qy | 576     | GG   | AAGACCTTTAAACCCAATTGCAGGCTCTGTGCCGCTGTGATGACGTGGCTTCA   | 635 |            |      |        |    |      |    |
| Dd | 302     | GGG  | GAGACCTTCCAGCCCCACTGCAGCATCCGCTGCCGCTCGGAGGACGGCGCTCAC  | 361 |            |      |        |    |      |    |
| Qy | 636     | TGC  | TGCGCTGTGCAGTGAGGATGTCGGCTGCCCAGCTGGGACTGCCACAGCCCCA    | 695 |            |      |        |    |      |    |
| Dd | 362     | TGC  | TGCGCTGTGCACGAGGATGTGCGGCTGCCAGCTGGGACTGCCCCACCCCA      | 421 |            |      |        |    |      |    |
| Qy | 696     | AGA  | TACAGTGGCCAGGAAAGTGTGCCCGCAGTGGGTATGTGACACGAGGAGTGACAC  | 755 |            |      |        |    |      |    |
| Dd | 422     | AGG  | TGAGTCTCTGGCAAGTGTGCTCCCTGAGTGGGTGTGCGGCCAAGAGGGGACTG   | 481 |            |      |        |    |      |    |
| Qy | 756     | GG   | ATCAGCGCTCCACGGCGCAGGACACAACACTTTCTGCCCTTTGTCACTTCCTCT  | 815 |            |      |        |    |      |    |
| Dd | 482     | GGG  | ACCCAGCCCTTCCAGCCCCAAGGACCCAGTTTTCTGGGCCTTGTCTTCTCTGCC  | 541 |            |      |        |    |      |    |
| Qy | 816     | GCT  | TGCTCTTGTCCAAATTTGAGACACAGCCTGGGGCCCTCTCAACCACTGTGG     | 875 |            |      |        |    |      |    |
| Dd | 542     | CTT  | GGTCTCCCTTCCCAGAATGAGGACCGGGCTGGGACCCCTCTCGACCACTGTGG   | 601 |            |      |        |    |      |    |
| Qy | 876     | CTGG | 880   |     |            |      |        |    |      |    |
| Dd | 602     | CTGG | 606   |     |            |      |        |    |      |    |

|            |  |             |            |              |                              |  |  |  |  |  |  |  |
|------------|--|-------------|------------|--------------|------------------------------|--|--|--|--|--|--|--|
| RESULT 5   | BM043988   | 750 bp      | mRNA       | linear       | EST 07-NOV-2001              |  |  |  |  |  |  |  |
| LOCUS      | BM043988   | 603620978F1 | NIH_MGC_40 | Homo sapiens | CDNA clone IMAGE:5446794 5', |  |  |  |  |  |  |  |
| DEFINITION | mRNA sequence.   |             |            |              |                              |  |  |  |  |  |  |  |
| ACCESSION  | BM043988   |             |            |              |                              |  |  |  |  |  |  |  |
| VERSION    | BM043988.1   | GI:16773255 |            |              |                              |  |  |  |  |  |  |  |
| KEYWORDS   | EST.   |             |            |              |                              |  |  |  |  |  |  |  |
| SOURCE     | human.   |             |            |              |                              |  |  |  |  |  |  |  |
| ORGANISM   | Homo sapiens   |             |            |              |                              |  |  |  |  |  |  |  |
| REFERENCE  | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. |             |            |              |                              |  |  |  |  |  |  |  |
| AUTHORS    | 1 (bases 1 to 750)   |             |            |              |                              |  |  |  |  |  |  |  |
| TITLE      | NTH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .  |             |            |              |                              |  |  |  |  |  |  |  |
| JOURNAL    | National Institutes of Health, Mammalian Gene Collection (MGC)   |             |            |              |                              |  |  |  |  |  |  |  |
| COMMENT    | Unpublished (1999)   |             |            |              |                              |  |  |  |  |  |  |  |
|            | Contact: Robert Strausberg, Ph.D.  |             |            |              |                              |  |  |  |  |  |  |  |
|            | Email: <a href="mailto:cgapbs-r@mail.nih.gov">cgapbs-r@mail.nih.gov</a>  |             |            |              |                              |  |  |  |  |  |  |  |
|            | Tissue Procurement: DCTD/DTP   |             |            |              |                              |  |  |  |  |  |  |  |
|            | cDNA Library Preparation: Ling Hong/Rubin Laboratory   |             |            |              |                              |  |  |  |  |  |  |  |
|            | cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  |             |            |              |                              |  |  |  |  |  |  |  |
|            | DNA Sequencing by: Incyte Genomics, Inc.   |             |            |              |                              |  |  |  |  |  |  |  |
|            | Clone distribution: MGC clone distribution information can be  |             |            |              |                              |  |  |  |  |  |  |  |
|            | found through the I.M.A.G.E. Consortium/LLNL at:   |             |            |              |                              |  |  |  |  |  |  |  |
|            | <a href="http://image.llnl.gov">http://image.llnl.gov</a>  |             |            |              |                              |  |  |  |  |  |  |  |
|            | Plate: L14CM1929   | row: k      | column: 19 |              |                              |  |  |  |  |  |  |  |

| FEATURES              | High quality sequence stop: 714.  |
|-----------------------|---|
| source                | Location/Qualifiers   |
|                       | 1. 750  |
|                       | /organism="Homo sapiens"  |
|                       | /db_xref="taxon:9606"   |
|                       | /clone="IMAGE:5446794"  |
|                       | /clone_lib="NIH_MGC_40"   |
|                       | /tissue_type="carcinoma, cell line"                                     |
|                       | /lab_host="DH10B (phage-resistant)"                                     |
|                       | /notes="Organ: prostate; Vector: pOTB7; Site:1: XhoI;                   |
|                       | Site 2: EcoRI; cDNA made by oligo-dT priming.                           |
|                       | Directionally cloned into EcoRI/XhoI sites using the                    |
|                       | following 5' adaptor: GCCACGAG(G). Library constructed by               |
|                       | Ling Hong in the laboratory of Gerald M. Rubin (University              |
|                       | of California, Berkeley) using ZAP-cDNA synthesis kit                   |
|                       | (Stratagene) and Superscript II RT (Life Technologies).                 |
|                       | Note: this is a NIH_MGC Library."                                       |
| BASE COUNT            | 112 a 267 c 246 g 125 t   |
| ORIGIN                |   |
| Query Match           | 45.4%; Score 400.6; DB 13; Length 750;                                  |
| Best Local Similarity | 79.5%; Pred. No. 6.6e-92;   |
| Matches               | 511; Conservative 0; Mismatches 129; Indels 3; Gaps 3;                  |
| Qy                    | 243 GGTGACATGAGGGCGACGCCACCTGATCCATCTTCTGGCCACTTCTCTCTGCTGCTGCTGCTC 302 |
| Db                    | 10 GGTGACATGAGGGCGACACCAAGAGCCACCTCTGGCTTCTTCCCTCTCTGCTGCTGCTC 69       |
| Qy                    | 303 TCAATGTGTGTGCCAGCTGTGCCGACACCCCTGTACCTGTCTTGGACACACACCCCCAG 362     |
| Db                    | 70 TCAAGGTGGTACCCAGCTGTGCCGACACCATGTACCTGCCCTGGCCACCTCCCGA 129          |
| Qy                    | 363 TGGCCACAGGGGTACCCCTGTGTGGATGTGGGTGTGGGTGTAAAGTGTGCACGG 422          |
| Db                    | 130 TGGCCCGTGGGATACCCCTGTGTGGATGTGGGTGTGGGTGTATGTGACCG 189              |
| Qy                    | 423 AGCTGGGGAGRCCTGCGACACCTGCATGTCTGGACCCCA-GCAGGGCCTGCTTG 481          |
| Db                    | 190 CGCTGGGGAGGCTGCGACCACTCCAGCTCTGGACCCATGCGAGGGCTGTGCTG 249           |
| Qy                    | 482 TCAGCTGGGGCAGGCCCTGGCGCCATGGGGCTGTGTGTCTTGGATGAGATGACGG 541         |
| Db                    | 250 CCAGCCGGGGCAGGCCCGTGGACGGGGGCCCTGTGCTCTTGGCAGAGACACAG 309           |
| Qy                    | 542 TACGTGTGAGGTGAATGGCCGCAAGTACCTGGATGAGAGACACCTTTAAACCCCAATTGCAG 601  |
| Db                    | 310 CAGCTGTGAGGTGAACGGCCGCTGTATCGGAAGGGAGACCTTCACGCCACCTGCAG 369        |
| Qy                    | 602 GGTCTGTGCCGTGTGTATCAGCGTGTCACTGCCTGCCCTGTGCAGTGCAGATGT 661          |
| Db                    | 370 CATCCGCTGCCGTGTGGAGGACGGCGCTTCACCTGGTGGCGCTGTGCAGCAGAGATGT 429      |
| Qy                    | 662 GCGGCTGCCAGTGGGACTGCCACGCCCCAAAGAAATACAGGTGCCAGGAAGTGCCTG 721       |
| Db                    | 430 GCGGCTGCCAGTGGGACTGCCCCACCCAGAGGGGTGAGGTCTTGGGCAAGTGCCTG 489        |
| Qy                    | 722 CCCCGAGTGGGTATGTGACCGAGGTGACACCGGGATGCC-AGCGCTCCACGGCCCAAG 780      |
| Db                    | 490 CCTGTAGTGGGTGTGGGCCAAGGAGGGGACTTGGGGACCCAAAGCCCCCTTCCAGCCCAAG 549   |
| Qy                    | 781 GACACCAACTTTTGGGCCCTTGTCACTCCTGCCCTGTCTGTATGCTCTTGTCTTCCAAATGGA 840 |
| Db                    | 550 GACCCCAAGTTTCTGGCCCTGTCTCTTCCCTGCCCTGTGTGCTGCTGCCAGATGGA 609        |
| Qy                    | 841 GCACAGCCTGGGCCCCCTGTCTCAACCACTGTGGGTGGGCAT 893                      |
| Db                    | 610 GCACGGCTGGGACCTGTGTCTGA-CACCTGTGGCTGGGCAT 651                       |

[illegible]

```

5', mRNA sequence.
ACCESSION BM921531
VERSION BM921531.1 GI:19371910
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgc.nci.nih.gov/
1 (bases 1 to 1006)
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12788 row: b column: 18
High quality sequence stop: 593.
Location/Qualifiers
1..1006
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5753009"
/clone_lib="NIH_MGC_115"
/lab_host="DH10B"
/Note="Organ: pooled brain, lung, testis; Vector:
pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."
BASE COUNT 148 a 368 c 317 g 169 t 4 Others
ORIGIN
Query Match 45.4%; Score 400.6; DB 14; Length 1006;
Best Local Similarity 78.7%; Pred. No. 7.6e-92;
Matches 489; Conservative 0; Mismatches 131; Indels 1; Gaps 1;
QY 256 GCAGCCCACTGATCCATCTTCGGCCACTTCTCTGCTGCTTCTCTCAATGGTGTG 315
DB 1 GCACACCGAAGACCCACCTCTGCGCTTCTCTGCTGCTTCTCTCAAGGTGCGTA 60
QY 316 CCAGCTGTGCGGACACCTCTAGCTGTCTTGGACACACCCAGTGCACAGGGGG 375
DB 61 CCCAGCTGTGCGGACACCATGACTGCTGCTGCGCCCTGCGCAGTCCCGCTGGAG 120
QY 376 TACCCTGTGTGGATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 435
DB 121 TACCCTGTGTGGATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
QY 436 CTTGCGACCACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 495
DB 181 CTTGCGACCACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
QY 496 GCCTGTGGGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 555
DB 241 GACCGGTGGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 300
QY 556 ATGGCGCGAGGTACCTGGATGAGAGACCTTTAAACCCCAATTTGAGGGTCTCTG 615
DB 301 ACGGCGCGCTGTATCGGAAGGGGAGACCTTCCAGCCCACTGACGATCCGCTCG 360
QY 616 GTGATGACGGTGGCTTACCTGCCCTGCGCTGTGAGTGTGAGGATGTGGGCTGCC 675

```

---

```

361 GCGAGGACGGGGGCTTACCTGCGTGGCGCTGTGAGCGAGGATGTGCGGCTGCCAGCT 420
QY 676 GGGACTGCCAGCCCAAGAAATACAGGTGCCAGAAAGTCTGCCCGAGTGGGTAT 735
DB 421 GGGACTGCCAGCCCAAGAGGAGGTCGAGGTCTGGCAAGTGTGCTGAGTGGGTGT 480
QY 736 GTGACGAGGAGTGACACCGGCGTCCAGCGCTCCAGCGCGGACGACCACTTCTG 795
DB 481 GCGGCAAGGAGGGGACTGGGGACCCAGCCCTTCCAGCCCAAGGACCCCAAGTTCTG 540
QY 796 CCCTGTGTCACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 854
DB 541 GCTTGTCTCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
QY 855 CCCTGCTCAACACCACTGTGGG 875
DB 601 CCCTGCTGACCAACCCCTGTG 621

RESULT 7
BI822142
LOCUS BI822142
DEFINITION BI822142.1 GI:15933692
ACCESSION BI822142
VERSION BI822142
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgc.nci.nih.gov/
1 (bases 1 to 886)
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Inyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11450 row: j column: 11
High quality sequence stop: 865.
Location/Qualifiers
1..886
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5180650"
/clone_lib="NIH_MGC_115"
/lab_host="DH10B"
/Note="Organ: pooled brain, lung, testis; Vector:
pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."
BASE COUNT 134 a 310 c 288 g 154 t
ORIGIN
Query Match 44.7%; Score 394.6; DB 13; Length 886;
Best Local Similarity 77.7%; Pred. No. 2.5e-90;
Matches 489; Conservative 0; Mismatches 139; Indels 1; Gaps 1;
QY 255 GGCACCCCACTGATCCATCTTCTGCGCCACTTCTCTGCTGCTGCTCTCTCAATGGTGT 314
DB 1 GGCACACCGAAGACCCACCTCTGCGCTTCTGCTGCTTCTCTCTCTCTCTCTCTCT 60

```





```

Db      382  TCGGAGGGGAGACCTTCCACCCCACTGACGATCCGCTGCCGCTGGAGGACGGGG 441
QY      629  CTTACCTGCTCCGCTGTGACGATGTGGGCTGCCAGCTGGGACTGCCACG 688
Db      442  CTTACCTGCTCCGCTGTGACGATGTGGGCTGCCAGCTGGGACTGCCACG 501
QY      689  CCCAAGAGATACAGGTGCGCAGGAAGTGTGCCCCGAGTGGGTATGTGACCGGGAT 748
Db      502  ACCAAGAGGTCGAGGTCCTGGCAAGTGTGCTCCCTGAGTGGTGTGCGGCAAGGAGG 561
QY      749  GACACGGGATCCAGCGCTCCACGGCGCAAGGACACCAACTTCTGCTGCTGCTCACT-- 806
Db      562  GGGACTGGGACCCAGACCTTACAGCCCAAGGACCCAGATCTGGCTTGTCTCTC 621
QY      807  CTTGCTGCTGCTGATGCTCTTGTCCAAATTGGAGCAGACGCTGGGG--CCCCCTGCTCAAC 865
Db      622  CCATGCCCACTGTGTCCCTGACCAAGTGGACGACGCTTGGGGAACCTGCTGCAC 681
QY      866  CACCTGTGGGCTGGG 880
Db      682  CACATGTGGGCTGG 696

RESULT 11
LOCUS   BI826781
DEFINITION 603077268f1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5169065 5',
mRNA sequence.
ACCESSION BI826781
VERSION   BI826781.1 GI:15938331
KEYWORDS EST.
SOURCE    human.
ORGANISM  Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS   Mammalia; Eutheraia; Primates; Catarrhini; Hominidae; Homo.
TITLE     NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL   National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT   Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11420 row: g column: 18
High quality sequence stop: 788.
Location/Qualifiers
1..800
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5169065"
/clone_lib="NIH_MGC_119"
/lab_host="DH10B"
/notes="Organ: brain; Vector: pCMV-SPORT6; Site:1: NotI;
Site:2: EcoRV (destroyed); RNA source normal medulla from
anonymous male age 27. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.3 kb, insert size range
0.9-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(invitrogen). Research Genetics tracking code 013. Note:
this is a NIH_MGC Library."

BASE COUNT 115 a 284 c 264 g 136 t 1 others
ORIGIN

Query Match 41.6%; Score 367.2; DB 13; Length 800;
Best Local Similarity 79.0%; Pred. No. 2.4e-83;
Matches 486; Conservative 0; Mismatches 124; Indels 5; Gaps 4;

```

```

QY      269  CCATCTTCTGGCACTTCTCTCTCAATGGTGTGTGCCAGCTGTGCCG 328
Db      14   CCACCTCTGGCTTCTCCCTCTCTCTCAAGGTGCGGTACCCAGCTGTGCC 73
QY      329  GACACCTGTACCTTCTCTTGGACACACCCAGTGGCCACAGGGGTGTGCCCTGGTGT 388
Db      74   GACACCATGTACGTCCCTTGGCCACCTCCCGATGCCGCTGGGAGTACCCCTGGTGT 133
QY      389  GGATGGCTGGGCTGTAAAGTGTGACAGGAGGTGGGGAGTCTCGGACCACT 448
Db      134  GGATGGCTGGGCTGTCCGGGTATGTGACGGGGCTGGGGAGACCTCTGGACCACT 193
QY      449  GCATCTCTGCGACCCAGCGGCTGTGTTGTGAGCTGGGGGAGCGCTGGCGGCA 508
Db      194  CCACGCTCTGCGAGCGCACCA--GGCCTGTGTCAGCCCGGGGAGGAGCCGGTGGCG 252
QY      509  TGGGGCTGTGTCTCTTGGATGAGGATGACGTTAGCTGTGAGTGAATGGCCGAGGTA 568
Db      253  GGGGCCCTGTGCCCTTGGCAGAGGACGACGAGCTGTGAGGTGAACGGCGCTGTA 312
QY      569  CTTGATGAGAGACCTTTAAACCAATTGAGGAGTCTGTGCCGCTGTGATGAGGTTG 628
Db      313  TCGGAAGGGAGACCTTCCAGCCCACTGACAGATCCGCTGCCGCTCGGAGGAGCGG 372
QY      629  CTTACCTGCTGCGCTGTGACGATGAGGATGTGCGGCTGCCAGCTGGGACTGCCACG 688
Db      373  CTTACCTGCTGCGCTGTGACGAGGATGTGCGGCTGCCAGCTGGGACTGCCACCA 432
QY      689  CCCNAGAGATACAGTGTGCGAGGAAAGTGTGCCCCAGTGGGTATGTGACCGAGGAGT 748
Db      433  CCCAGAGAGGTGCGAGGTCCTGGGCAAGTGTGCCCTGAGTGGGTGTGCG--GCAAGGAG 491
QY      749  GACACCGGCGATCCAGCGCTCCACGGCGCAAGGACACCAACTTCTGCTGCTGTACATCC 808
Db      492  GGGACTGGGACCCAGCCCTTCCAGCCCAAGGACCCAGTTCCTGCGCTTGTCTCTTC 551
QY      809  TGCCTCTGATGCTCTTGTCCAAATTTGGAGCACACCTGGGCGCCCTGCTCAACAC 868
Db      552  -CCTGCCCCCTGTGCTCCCTGCCAGATGGAGNACAGGC--TGGGACCTGCTCGACCA 608
QY      869  CTGTGGGCTGGCAT 883
Db      609  CTGTGGGCTGGCAT 623

```

```

RESULT 12
LOCUS   BI825652
DEFINITION 603072631f1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5164614 5',
mRNA sequence.
ACCESSION BI825652
VERSION   BI825652.1 GI:15937202
KEYWORDS EST.
SOURCE    human.
ORGANISM  Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS   Mammalia; Eutheraia; Primates; Catarrhini; Hominidae; Homo.
TITLE     NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL   National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT   Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11408 row: n column: 07
High quality sequence stop: 877.

```



| FEATURES              | Location/Qualifiers  |
|-----------------------|--|
| source                | 1..888   |
|                       | /organism="Homo sapiens"   |
|                       | /db_xref="taxon:9606"  |
|                       | /clone="IMAGE:5164614"   |
|                       | /clone_lib="NIH_MGC_119"   |
|                       | /tissue_type="medulla"   |
|                       | /lab_host="DH10B"  |
|                       | /notes="Organ: brain; Vector: pCMV-SPORT6; site_1: NotI;<br>site_2: EcoRV (destroyed); RNA source normal medulla from<br>anonymous male age 27. Library is oligo-dT primed and<br>directionally cloned (EcoRV site is destroyed upon<br>cloning).. Average insert size 1.3 kb, insert size range<br>0-9.3 kb. Library is normalized and enriched for<br>full-length clones and was constructed by C. Gruber<br>(Invitrogen). Research Genetics tracking code 013. Note:<br>this is a NIH_MGC Library." |
| BASE COUNT            | 130 a 308 c 288 g 162 t  |
| ORIGIN                |  |
| Query Match           | 41.6%; Score 367.2; DB 13; Length 888;   |
| Best Local Similarity | 77.1%; Pred. No. 2.5e-83;  |
| Matches               | 472; Conservative 0; Mismatches 138; Indels 2; Gaps 2;   |
| QY                    | 273 CTCTGTGCCACTTCCTCCTCCCTCTGCTCTCAATGGTGTCGCCAGCTGTGCCGACA 332   |
| Db                    |  |
|                       | 15 CCTCTGGCTTCTCCCCTCTGCTCTCTCAAAGTGCCTACCCAGCTGTGCCGACA 74  |
| QY                    | 333 CCCTGTACCTGTCTTTGGACACCACCCAGTGCACCACAGGGGGTACCCCTGGTGTGGAT 392  |
| Db                    |  |
|                       | 75 CCATGTACCTTGCCTCGCCCTGCCACCTCCCCGATGCCCGCTGGAGTAGCCCTGGTGTGGAT 134  |
| QY                    | 393 GGCTGTGGCTGCTTAAAGTGTFGCACGAGGCTGGGGAGTCTCTCGGACCACCTGCAT 452  |
| Db                    |  |
|                       | 135 GGCTGTGGCTGCTGCCGGGTATGTGACGGGGCTGGGGAGCCCTCGGACCACTCCAC 194   |
| QY                    | 453 GTCTGCGACCCACGACGGCCCTGGTTGTTCAGCTGGGGCAGGCCCTGGCGGCATGEGG 512   |
| Db                    |  |
|                       | 195 GTCTGCGACCCACGACGGCTGTGTCTGCCAGCCCGGGCAGGACCCGGTGGCCGGGG 254   |
| QY                    | 513 GCTGTGTGCTCTTGGATGAGGTAGCGGTAGCTGTGAGGTGAATGCCGCGAGGTACTGT 572   |
| Db                    |  |
|                       | 255 GCCTGTGCTCTTGGCAGAGGACGACAGCAGCTGTGAGGTGAACGCGCCCTGTATCGG 314  |
| QY                    | 573 GATGGAGAGACCTTTAAACCCAATTGCAGGGTCTCTGCGCGCTGTGATCAGGTGGCTTC 632  |
| Db                    |  |
|                       | 315 GAAGGGGAGACCTTCAGCCCCACTGCAGCATCCGCTGCCGCTGCGAGGACGGCGGCTC 374   |
| QY                    | 633 ACCTGCGTCCGCTGTGCTAGTAGGATGTGCGGCTGCCAGCTGGGACTGCCACAGCCCC 692   |
| Db                    |  |
|                       | 375 ACCTGCGTCCGCTGTGCACGAGGATGTGCGGCTGCCAGTCCAGCTGGGACTGCCCCACCCC 434  |
| QY                    | 693 AAGAGAATACAGTGCACGAGAAAGTGTGCCCGAGTGGGTATGTGACCAAGGAGTAGTGA 752  |
| Db                    |  |
|                       | 435 AGGAGGGTCAGGTCTCTGGCAAGTGTCTGCCCTAGTGGGTGTGCGGCAAGGAGGGGA 494  |
| QY                    | 753 CCGCGCATCAGCGCTCCACGGCGCAGGACACCACTTTCGCCCTTGTGCAC-TCTGTC 811  |
| Db                    |  |
|                       | 495 CTGGGACCCAGCCCCCTTCAGGCCAAGGACCCAGTTTCTGGCCCTTGTCTTTCCTG 554   |
| QY                    | 812 CTCGTGTGATGCTCTTGTGTCAAATTGGAGCACAGCTGGGGGCCCTGTCTCAACCACTG 871  |
| Db                    |  |
|                       | 555 GCCCCCTGGTGTCCCTGCCCAAGATGGAGCAC-GGCTGTGGACCTGTCTCGACCACTG 613   |
| QY                    | 872 TGGGCTGGGCAT 883   |
| Db                    |  |
|                       | 614 TGGGCTGGGCAT 625   |
| RESULT 13             |  |
| BG928868              |  |
| LOCUS                 | BG928868 620 bp mRNA linear EST 06-NOV-2001  |
| DEFINITION            | HNC57-1-D9.R HNC (Human Normal Cartilage) Homo sapiens cDNA, mRNA  |

```

QY 724 CCGAGTGGGTATGTGACACAGGAGTGACACCGCGGATCCAGCGCTCCAGCGGCGCAAGGAC 783
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 559 CTGAGTGGGTGTGCGGCCAAGGAGGGGACTGGGGACCCAGCCCTTCAGCCCAAGGAC 618
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 784 AC 785
Db 619 CC 620

RESULT 14
BI457367
LOCUS 916 bp mRNA linear EST 21-AUG-2001
DEFINITION 60318589F1 NTH_MGC_42 Homo sapiens cDNA clone IMAGE:5258398 5',
mRNA sequence.
ACCESSION BI457367
VERSION BI457367.1 GI:15248023
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 916)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1870 row: a column: 23
High quality sequence stop: 778.
FEATURES
    source
        1..916
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:5258398"
            /clone_lib="NIH_MGC_42"
            /tissue_type="epithelioid carcinoma cell line"
            /lab_host="DH10B (phage-resistant)"
            /note="Organ: pancreas; Vector: pOMB7; Site:1: XhoI;
            Site:2: EcoRI; cDNA made by oligo-dT priming.
            Directionally cloned into EcoRI/XhoI sites using the
            following 5' adaptor: GGCACGAG(G). Size-selected >500bp
            for average insert size 1.8kb. Library constructed by Ling
            Hong in the laboratory of Gerald M. Rubin (University of
            California, Berkeley) using ZAP-cDNA synthesis kit
            (Stratagene) and Superscript II RT (Life Technologies).
            Note: this is a NIH-MGC Library."
BASE COUNT 136 a 305 c 304 g 170 t 1 others
ORIGIN
Query Match 41.2%; Score 363.4; DB 13; Length 916;
Best Local Similarity 79.2%; Pred. No. 2.4e-82;
Matches 480; Conservative 0; Mismatches 122; Indels 4; Gaps 4;

QY 279 GCACCTTCCTGCTGCTCTCTCAATGGTGTGCGCCAGCTGTGCGGACACCCCTGT 338
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 GCCTTCCTCCTCTGCTGCTCTCTCAAGGTGCGTACCGAGCTGTGCGCCGACACCATGT 60
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 339 ACCTCTCTCTGACACACCCAGTCCACAGGGGGTACCCCTGCTGCTGGATGGCTGT 398
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 ACCTGCCCTGG-CACCTCCCGATGCCCGCTGGGAGTACCCCTGGTGTGATGGCTGT 119
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 399 GCGTCTCTAAAGTGTGTGCACGGAGGTGGGGAGTCTCTGCGACCACTCATCTCTGC 458
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 120 GCGTCTCGGGTATGTGCACGGGGGCTGGGGAGCCCTGGCACCCTCCACCTCTGCG 179
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 459 GACCCACGACGGGCTGTTGTGACGCTGGGGCAGGCCCTGGGGCCATGGGGCTGTG 518
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 180 GACGCCAGCCA-GGCTGTGCTCTGCCAGCCCGGGCAGGACCCCGTGGACGGGGGCCCTG 238
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 519 TGTCTCTTGGATGAGGATGACGGTAGCTGTGAGGTGAATGGCCGAGGTACCTCGATGGA 578
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 239 TGCCTCTTTGGCAGAGGACGACAGCAGCTGTGAGGTGAACGGCCGCTGTATCGGGAAGG 298
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 579 GAGACCTTTAAACCAATTCAGAGGTCTGTGCCCTGTGTATGACGGTGGCTTACCTTGC 638
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 299 GAGACCTTCCAGCCCACTGCAGCATCCGCTGCCCTGGAGGAGGGGGCTTACCTTGC 358
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 639 CTGCCGCTGTGAGTGTGAGGATGTGCGGTGCCAGCTGGGAGTGCACGCCCCCAAGAGA 698
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 359 CTGCCGCTGTGAGGAGGATGTGCGGTGCCAGCTGGGAGTGCACGCCCCCAAGAGA 418
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 699 ATACAGGTGCCAGGAAGTGTGCCCGGAGTGGGTATGTGACCAAGGAGTGACACCGCG 758
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 419 GTCGAGGTCTGGGCAAGTGTGCTGAGTGGGTGTGCGGCCAAGGGGAGTGGG 478
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 759 ATCCAGCGCTCCACGCGCAAGGACACCAACTTTCTGCCCTTGTCACTCCTGCCCTGTCT 818
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 479 ACCCAGGCCCTTCACAGCCCAAGGACCCAGTTTCTGGCTTGTCTCTTCCCTGCCCTCT 538
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 819 G-ATGCTCTCTGTGCCAAATTTGGAGCACAGCCTGGGGCCCTGCTCAACCACTGTGGCT 877
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 539 GTGTGTCCCTGCCCAAGATGGAGCACGCGCTGGGGACCTGCTCGA-CACCTGTGGCT 597
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 878 GGGCAT 883
Db 598 GGGCAT 603

RESULT 15
BM543799
LOCUS 1166 bp mRNA linear EST 20-FEB-2002
DEFINITION AGENCOURT_6492297 NIH_MGC_125 Homo sapiens cDNA clone IMAGE:5589134
5', mRNA sequence.
ACCESSION BM543799
VERSION BM543799.1 GI:18774503
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1166)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Invitrogen
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LAM12361 row: f column: 15
High quality sequence stop: 547.
FEATURES
    source
        1..1166
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:5589134"
            /clone_lib="NIH_MGC_125"
            /lab_host="DH10B"
            /note="Organ: ovary (pool of 3); Vector: pCMV-SPORT6;
            Site:1: EcoRV (destroyed); Site:2: NotI; RNA source pool
            of three ovaries, from females ranging in age from 38 to
            49 yo. Library is oligo-dT primed and directionally cloned
            (EcoRV site is destroyed upon cloning). Average insert
            size 2.1 kb, insert size range 1-3.5 kb. Library is
            normalized and enriched for full-length clones and was

```

Search completed: July 29, 2003, 02:51:26  
Job time : 1569.96 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 1, 2003, 01:19:19 ; Search time 2638 Seconds  
(without alignments)  
2758.037 Million cell updates/sec

Title: US-10-010-408-2

Perfect score: 1440

Sequence: 1 MRGSLPLHLATSLFLCLLSM.....LCLPRFLAARSHSWNSAF 250

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh  
-Q/cgn2\_1/USPTO\_spool/US10010408/runat\_25072003\_101855\_17924/app\_query.fasta\_1.391  
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=BITS -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THRESHOLD=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10010408 -CGN\_1\_1\_2496 -runat\_25072003\_101855\_17924 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_hgt.\*

3: gb\_in.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vi.\*

15: em\_ba.\*

16: em\_fun.\*

17: em\_hum.\*

18: em\_in.\*

19: em\_mu.\*

20: em\_om.\*

21: em\_or.\*

22: em\_ov.\*

23: em\_pat.\*

24: em\_ph.\*

25: em\_pl.\*

26: em\_ro.\*

27: em\_sts.\*

28: em\_un.\*

29: em\_vi.\*  
30: em\_htg\_hum.\*  
31: em\_htg\_inv.\*  
32: em\_htg\_other.\*  
33: em\_htg\_mus.\*  
34: em\_htg\_pln.\*  
35: em\_htg\_rod.\*  
36: em\_htg\_mam.\*  
37: em\_htg\_vrt.\*  
38: em\_sy.\*  
39: em\_htgo\_hum.\*  
40: em\_htgo\_mus.\*  
41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score  | Query Match % | Length | DB ID | Description          |
|------------|--------|---------------|--------|-------|----------------------|
| 1          | 1390.5 | 96.5          | 1741   | 10    | AF259981 Rattus no   |
| 2          | 1308.5 | 90.9          | 1734   | 6     | AR210324 Sequence    |
| c 3        | 1308.5 | 90.9          | 1734   | 6     | AR210325 Sequence    |
| 4          | 1308.5 | 90.9          | 1734   | 10    | AF100778 Mus muscu   |
| 5          | 1308.5 | 90.9          | 1739   | 10    | AF126063 Mus muscu   |
| 6          | 1064   | 73.9          | 1266   | 6     | AX076919 Sequence    |
| 7          | 1064   | 73.9          | 1266   | 6     | AX464186 Sequence    |
| 8          | 1064   | 73.9          | 1283   | 9     | AF083500 Homo sapi   |
| c 9        | 1064   | 73.9          | 1293   | 6     | AR210322 Sequence    |
| 10         | 1064   | 73.9          | 1293   | 6     | AR210323 Sequence    |
| 11         | 1064   | 73.9          | 1309   | 9     | AF074604 Homo sapi   |
| 12         | 1064   | 73.9          | 1427   | 9     | AF100780 Homo sapi   |
| 13         | 1064   | 73.9          | 1450   | 9     | BC017782 Homo sapi   |
| 14         | 1047   | 72.7          | 738    | 6     | AR210337 Sequence    |
| 15         | 1045.5 | 72.6          | 841    | 6     | AR210338 Sequence    |
| 16         | 575.5  | 40.0          | 1585   | 5     | NV1271167 Notoptha   |
| 17         | 564    | 39.2          | 2075   | 6     | AR018957 Sequence    |
| 18         | 564    | 39.2          | 2075   | 6     | AR118595 Sequence    |
| 19         | 564    | 39.2          | 2075   | 6     | AR130891 Sequence    |
| 20         | 564    | 39.2          | 2075   | 6     | AX206708 Sequence    |
| 21         | 564    | 39.2          | 2075   | 6     | I11636 Sequence 1    |
| 22         | 564    | 39.2          | 2075   | 6     | I32210 Sequence 1    |
| 23         | 564    | 39.2          | 2075   | 9     | HUMCONGRO            |
| 24         | 564    | 39.2          | 2312   | 9     | HSC7GF               |
| 25         | 564    | 39.2          | 2998   | 6     | AR119211 Sequence    |
| 26         | 564    | 39.2          | 2998   | 6     | AR151276 Sequence    |
| 27         | 553    | 38.4          | 2190   | 5     | AF463517 Gallus ga   |
| 28         | 553    | 38.4          | 2288   | 5     | GA298335 Gallus ga   |
| 29         | 550.5  | 38.2          | 1953   | 5     | XLU43524 Xenopus lae |
| 30         | 550.5  | 38.2          | 2267   | 6     | AX206706 Sequence    |
| 31         | 550.5  | 38.2          | 2267   | 10    | MUSFIS12B            |
| 32         | 550.5  | 38.2          | 2330   | 10    | MUSTGFB              |
| 33         | 550.5  | 38.2          | 2334   | 10    | BC006783 Mus muscu   |
| 34         | 550.5  | 38.2          | 2345   | 6     | AX402019 Sequence    |
| 35         | 550.5  | 38.2          | 2345   | 10    | AF120275 Rattus no   |
| 36         | 548.5  | 38.1          | 2338   | 6     | E37595 Monoclonal    |
| 37         | 548.5  | 38.1          | 2338   | 10    | AB023068             |
| 38         | 542.5  | 37.7          | 2350   | 6     | AR194011 Sequence    |
| 39         | 542.5  | 37.7          | 2350   | 6     | AR201286 Sequence    |
| 40         | 541.5  | 37.6          | 1975   | 6     | A28405 Sequence I    |
| 41         | 541.5  | 37.6          | 1975   | 6     | A28444 nov mRNA se   |
| 42         | 541.5  | 37.6          | 1976   | 5     | GGNOVMRNA            |
| 43         | 541    | 37.6          | 1805   | 5     | CHKCEP               |
| c 44       | 537    | 37.3          | 137964 | 2     | AC126895             |
| 45         | 533    | 37.0          | 1496   | 4     | SSU83916             |

ALIGNMENTS

RESULT 1

|   |  |  |                            |
|---|--|--|----------------------------|
| Db  | 502  | GGGGCAGGCCCTGGCGGCCATGGGGCTGTGTCTTGGATGAGATGACGGTGACTGT      | 561                        |
| Qy  | 101  | GluValAsnGlyArgArgTyrLeuAspGlyGluThrPheLysProAsnCysArgValLeu | 120                        |
| Db  | 562  | GAGTGAATGGCGCGCAGGTACCTGGATGGAGACCTTTAACCCCAATTGCAGGGTCCGTG  | 621                        |
| Qy  | 121  | CysArgCysAspGlyGlyPheThrCysLeuProLeuCysSerGluAspValArgLeu    | 140                        |
| Db  | 622  | TGCGCGTGTGATGACGGTGGCTTACCTGCCTCCCGCTGTGCAGTGAGGATGTGACGCTG  | 681                        |
| Qy  | 141  | ProSerTrpAspCysProArgProLysArgIleGlnValProGlyLysCysCysProGlu | 160                        |
| Db  | 682  | CCCAGCTGGACATGCCACGCCCAAGAGAATAACAGGTGCCAGAAAGTCTGCCCCGAG    | 741                        |
| Qy  | 161  | TrpValCysAspGlnGlyValThrProAlaIleGlnArgSerThrAlaGlnGlyHisGln | 180                        |
| Db  | 742  | TGGGTATGTACAGGAGGTGACACCGGGGATCCAGCGCTCCGCGCGGAAGACACCAA     | 801                        |
| Qy  | 181  | LeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThrAla | 200                        |
| Db  | 802  | CTTTCGTGCCCTGTCTACTCTCTGTGTGTCTCTTGGCCAAATTTGGAGCAGGCC       | 861                        |
| Qy  | 201  | TrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGlnAsn | 220                        |
| Db  | 862  | TGGGGCCCTGCTCAACACCTCTGGGCTGGGCATAGCCACCGAGTGTCCAACCAAGAC    | 921                        |
| Qy  | 221  | ArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProArgProCysLeuAlaAla | 240                        |
| Db  | 922  | CGATTCTGCCAATGGAGATCCAAACCGCGCTGTCTGCCACAGACCTTCCTGGCAGCC    | 981                        |
| Qy  | 241  | ArgSerHisSerSerTrpAsnSerAlaPhe                               | 250                        |
| Db  | 982  | AGAGGCCACAGCTCATGGACAGTGTCTTC                                | 1011                       |
| RESULT 2                                    |  |  |                            |
| LOCUS                                       | AR210324   | 1734 bp  | DNA linear PAT 20-JUN-2002 |
| DEFINITION                                  | Sequence 17 from patent US 6387657.  |  |                            |
| ACCESSION                                   | AR210324   |  |                            |
| VERSION                                     | AR210324.1   | GI:21512525  |                            |
| KEYWORDS                                    | Unknown.   |  |                            |
| SOURCE                                      | Unknown.   |  |                            |
| ORGANISM                                    | Unclassified.  |  |                            |
| REFERENCE                                   | 1 (bases 1 to 1734)  |  |                            |
| AUTHORS                                     | Botstein,D.A., Cohen,R.L., Goddard,A.D., Gurney,A.L., Hillan,K.J., Lawrence,D.A., Levine,A.J., Pennica,D., Roy,M. Ann. and Wood,W.I. |  |                            |
| TITLE                                       | WSP polypeptides and nucleic acids encoding same   |  |                            |
| JOURNAL                                     | Patent: US 6387657-A 17 14-MAY-2002;   |  |                            |
| FEATURES                                    | Location/Qualifiers  |  |                            |
| source                                      | 1..1734  |  |                            |
| BASE COUNT                                  | 355 a 491 c 495 g  | 393 t  |                            |
| ORIGIN                                      | /organism="unknown"  |  |                            |
| Alignment Scores:                           |  |  |                            |
| Pred. No.:                                  | 3,97e-67   | Length:  | 1734                       |
| Score:                                      | 1308.50  | Matches:   | 226                        |
| Percent Similarity:                         | 93.63%   | Conservative:  | 9                          |
| Best Local Similarity:                      | 90.04%   | Mismatches:  | 15                         |
| Query Match:                                | 90.87%   | Indels:  | 1                          |
| DB:   | 6  | Gaps:  | 1                          |
| US-10-010-408-2 (1-250) x AR210324 (1-1734) |  |  |                            |
| Qy  | 1  | MetArgGlySerProLeuIleHisLeuLeuAlaThrSerPheLeuCysLeuLeuSerMet | 20                         |
| Db  | 257  | ATGAGGGCAACCACTGATCCATCTCTGGCCATTCCTCTCTGCATTCCTCTCAATG      | 316                        |
| Qy  | 21   | ValCysAlaGlnLeuCysArgThrProCysThrCysProThrProGlnCysPro       | 40                         |
| Db  | 317  | GTGATTCCAGGTGGCCAGCACCCCTGTCCCTTGGACACACCCAGTCCCA            | 376                        |







```

/protein_id="AAD18058.1"
/db_xref="GI:4337060"
/translation="MRGNPLHLALLAIFSLIILMSQLCPACPCWTPPPQCPGVP
LVLDGCGRCVARRUGESDHLHVCDDPSQGLVCPQAGPSGRGAVCLFEEDDGSCEV
NGRRYLDGTFEPFNCVRLCRDDGGFTCLPLCSEDVRLSWDCPPRRRLQVPRGCCPE
WYCDQVMQAPTPQSSQAQGHQSALVTPASADGCPNPWSTAGPCSTTCGLGIAIATRS
NNRRFCQLEIQRRLCSFCLASRSHGSNSAF"
375 a 489 c 489 q 395 t
BASE COUNT

```

| Alignment Scores:                           |          |   |      |
|---|----------|---|------|
| Pred. No.:                                  | 3,98e-67 | Length:   | 1739 |
| Score:                                      | 1308.50  | Matches:  | 226  |
| Percent Similarity:                         | 93.63%   | Conservative:   | 9    |
| Best Local Similarity:                      | 90.04%   | Mismatches:   | 15   |
| Query Match:                                | 90.87%   | Indels:   | 1    |
| DB:   | 10       | Gaps:   | 1    |
| US-10-010-408-2 (1-250) x AF126063 (1-1739) |          |   |      |
| QY  | 1        | MetArgGlySerProLeuIleHisLeuLeuAlaThrSerPheLeuCysLeuSerMet       | 20   |
| DB  | 242      | ATGAGGGGCAACCACTGATCCATCTCTGGCCATTCTCTCTCTCATCTCTCAATG          | 301  |
| QY  | 21       | ValCysAlaGlnLeuCysArgThrProCysThrCysCysProTrpThrProGlnCysPro    | 40   |
| DB  | 302      | GTGTAITCCAGCTGTGCCAGCACCTGTGCGCTGTCTTGGACACACCCCAAGTCCCA        | 361  |
| QY  | 41       | GlnGlyValProLeuValLeuAspGlyCysGlyCysCysLysValCysAlaArgArgLeu    | 60   |
| DB  | 362      | CCGGGGGTACCCCTGGTGTGGATGGTGTGGCTGTGCTGTGCTGTGCTGCCGAGGCTG       | 421  |
| QY  | 61       | GlyGluSerCysAspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnPro    | 80   |
| DB  | 422      | GGGAGTCTTGCACCACTGATGCTGTGGACCCAGCCAGGCGCTGTTGTCAAGCT           | 481  |
| QY  | 81       | GlyAlaGlyProGlyGlyHisGlyAlaValCysLeuLeuAspGluAspGlySerCys       | 100  |
| DB  | 482      | GGGGCAGGCCCCAGTGGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT         | 541  |
| QY  | 101      | GluValAsnGlyArgArgTyrLeuAspGlyGluThrPheLysProAsnCysArgValLeu    | 120  |
| DB  | 542      | GAGGTGAAGCGCGCAGGTACCTGGATGGGAGACCTTTAAACCCCAATTGCAGGGTTTG      | 601  |
| QY  | 121      | CysArgCysAspGlyGlyPheThrCysLeuProLeuCysSerGluAspValArgLeu       | 140  |
| DB  | 602      | TGCGCGTGTGATACGGTGGTTTTCACCTGCTGCCGCTGTGCAGTGGAGTGTGCGGCTG      | 661  |
| QY  | 141      | ProSerTrpAspCysProArgProLysArgGlyIleGlnValProGlyLysCysCysProGlu | 160  |
| DB  | 662      | CCAGCTGGGACTGCCACGCCCGCCAGGAGNATACAGGTGCCAGAGAGGTGCTGCCCGAG     | 721  |
| QY  | 161      | TrpValCysAspGlnGlyVal---ThrProAlaIleGlnArgSerThrAlaGlnGlyHis    | 179  |
| DB  | 722      | TGGGTGTGTGACCAAGCAGTGATGCAGCGCGCAATCCAGCCCTCTCAGCCCCAAGGACAC    | 781  |
| QY  | 180      | GlnLeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThr    | 199  |
| DB  | 782      | CAACTTTTCTGCCCTTGTCACTTCTGCACTTCCGCGATGGCCCTGTCCAACTGGAGCACA    | 841  |
| QY  | 200      | AlaTrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGln    | 219  |
| DB  | 842      | GCTTGGGGCCCTGTCTACCAACCTGTGGGTGGGCATAGCCACCCGAGTATCCCAACCAG     | 901  |
| QY  | 220      | AsnArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProArgProCysLeuAla    | 239  |
| DB  | 902      | AACCGATTTCGCAACTGGAGATCCAGCGTTCGCTGTGTCTGTCTGTCCAGACCTGCCTGGCA  | 961  |
| QY  | 240      | AlaArgSerHisSerSerTrpAsnSerAlaPhe                               | 250  |
| DB  | 962      | TCCAGGAGCCAGCGCTCATGGAACTGGCTTC                                 | 994  |

## RESULT 6

|   |  |  |         |       |        |                 |
|---|--|--|---------|-------|--------|-----------------|
| AX076919                                    | AX076919   | Sequence 31 from Patent WO0105836.                           | 1266 bp | DNA   | linear | PAT 22-FEB-2001 |
| LOCUS                                       | AX076919   | Sequence 31 from Patent WO0105836.                           |         |       |        |                 |
| DEFINITION                                  | AX076919   | Sequence 31 from Patent WO0105836.                           |         |       |        |                 |
| ACCESSION                                   | AX076919   | Sequence 31 from Patent WO0105836.                           |         |       |        |                 |
| VERSION                                     | AX076919.1   | GI:13121575  |         |       |        |                 |
| KEYWORDS                                    | human.   |  |         |       |        |                 |
| SOURCE                                      | human.   |  |         |       |        |                 |
| ORGANISM                                    | Homo sapiens   |  |         |       |        |                 |
| REFERENCE                                   | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. |  |         |       |        |                 |
| AUTHORS                                     | Botstein,D., Goddard,A., Gurney,A.L., Hillan,K.J., Roy,M.A. and Wood,W.I.  |  |         |       |        |                 |
| TITLE                                       | Polypeptidic compositions and methods for the treatment of tumors  |  |         |       |        |                 |
| JOURNAL                                     | Patent: WO 0105836-A 31 25-JAN-2001;   |  |         |       |        |                 |
| FEATURES                                    | Genentech, Inc. (US)   |  |         |       |        |                 |
| source                                      | Location/Qualifiers  |  |         |       |        |                 |
|   | 1..1266  |  |         |       |        |                 |
|   | /organism="Homo sapiens"   |  |         |       |        |                 |
|   | /db_xref="taxon:9606"  |  |         |       |        |                 |
| BASE COUNT                                  | 216 a  | 418 c  | 390 g   | 242 t |        |                 |
| ORIGIN                                      |  |  |         |       |        |                 |
| Alignment Scores:                           |  |  |         |       |        |                 |
| Pred. No.:                                  | 3.56e-53   |  |         |       |        | 1266            |
| Score:                                      | 1064.00  |  |         |       |        | 184             |
| Percent Similarity:                         | 80.40%   |  |         |       |        | 17              |
| Best Local Similarity:                      | 73.60%   |  |         |       |        | 49              |
| Query Match:                                | 73.89%   |  |         |       |        | 0               |
| DB:   | 6  |  |         |       |        | 0               |
| US-10-010-408-2 (1-250) x AX076919 (1-1266) |  |  |         |       |        |                 |
| Qy  | 1  | MetArgGlySerProLeuIleHisLeuLeuAlaThrSerPheLeuCysLeuLeuSerMet | 20      |       |        |                 |
| Db  | 10   | ATGAGGACACACCGAAGACCCACCTCTGGCCCTTCTCCCTCTCTCTCTCTCTCAAG     | 69      |       |        |                 |
| Qy  | 21   | ValCysAlaGlnLeuCysArgThrProCysThrCysProThrThrProProGlnCysPro | 40      |       |        |                 |
| Db  | 70   | GTGCGTACCAGCTGTGCCCGACACCATGTACCTGCCCTTGGCCACCTCCCGATGCCG    | 129     |       |        |                 |
| Qy  | 41   | GlnGlyValProLeuValLeuAspGlyCysGlyCysCysValCysAlaArgArgLeu    | 60      |       |        |                 |
| Db  | 130  | CTGGAGTACCCCTGGTGTCTGGATGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGG     | 189     |       |        |                 |
| Qy  | 61   | GlyGluSerCysAspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnPro | 80      |       |        |                 |
| Db  | 190  | GGGAGGCCCTGGACCAACTCCAGCTCTGCAGGCCAGCCAGGSCCTGGTCTGCCAGCC    | 249     |       |        |                 |
| Qy  | 81   | GlyAlaGlyProGlyGlyHisGlyAlaValCysLeuLeuAspGluAspAspGlySerCys | 100     |       |        |                 |
| Db  | 250  | GGGCGAGGACCCGGTGGCGGGGCCCTGTGCTTGGCAGAGGACGACACGACGTGT       | 309     |       |        |                 |
| Qy  | 101  | GluValAsnGlyArgArgTyrLeuAspGlyGluThrPheLysProAsnCysArgValLeu | 120     |       |        |                 |
| Db  | 310  | GAGGTGAACGGCCGCTGTATCGGGAAGGGAGACCTTCCAGCCCCACTGCAGCATCCG    | 369     |       |        |                 |
| Qy  | 121  | CysArgCysAspGlyGlyPheThrCysLeuProLeuCysSerGluAspValArgLeu    | 140     |       |        |                 |
| Db  | 370  | TGCCGCTGCAGGACGGCGCTTACCTGCTGCGCTGTGCAGCAGGATGTGCGGCTG       | 429     |       |        |                 |
| Qy  | 141  | ProSerTrpAspCysProArgProLysArgIleGlnValProGlyLysCysProGlu    | 160     |       |        |                 |
| Db  | 430  | CCGACGTGGGACTGCCCCCAGGAGGGGTGCAGGTCTGGGCAAGTGTGCCCCGTGAG     | 489     |       |        |                 |
| Qy  | 161  | TrpValCysAspGlnGlyValThrProAlaIleGlnArgSerThrAlaGlnGlyHisGln | 180     |       |        |                 |
| Db  | 490  | TGGGTGTGGCCCAAGGAGGGGGGACTGGGACCCAGCCCTTCCAGCCCAAGNACCCAG    | 549     |       |        |                 |
| Qy  | 181  | LeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThrAla | 200     |       |        |                 |
| Db  | 550  | TTTTCTGGCCTGTCTCTCCCTGCCCTTGGTGTCTCCCTGCCAGAAATGGAGCAGGCC    | 609     |       |        |                 |

QY 201 TrpGlyProCysSerThrThrCysGlyLeuGlyLeuAlaThrArgValSerAsnGlnAsn 220  
 Db 610 TGGGACCCCTGCTGCACACCTGTGGTGGGCATGGCCACCCGGGTGTCCAACAGAAC 669

QY 221 ArgPheCysGlnLeuGluLeuGlnArgLeuCysLeuProArgProCysLeuAlaAla 240  
 Db 670 GCCTTCTCCGACTGGAGACCAGCCGCGCTGTGCCTGTCTCCAGGCGCTGCCACCCCTCC 729

QY 241 ArgSerHisSerSerTrpAsnSerAlaPhe 250  
 Db 730 AGGGTCCGACGTCCACAAACAGTGCCTTC 759

RESULT 7  
 AX464186  
 LOCUS  
 DEFINITION Sequence 319 from Patent WO0140466.  
 ACCESSION AX464186  
 VERSION AX464186.1 GI:21899109  
 KEYWORDS  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
 AUTHORS Baker,K.P., Beresini,M., Deforge,L., Desnoyers,L., Filvaroff,E.,  
 Gao,W.Q., Gerritsen,M.E., Goddard,A., Godowski,P.J., Gurney,A.L.,  
 Sherwood,S., Smith,V., Stewart,T.A., Tumas,D., Watanabe,C.K.,  
 Wood,W.L. and Zhang,Z.  
 TITL Secreted and transmembrane polypeptides and nucleic acids encoding  
 same  
 JOURNAL Patent: WO 0140466-A 319 07-JUN-2001;  
 Genentech Inc. (US)  
 FEATURES  
 source Location/Qualifiers  
 1. 1266  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"

BASE COUNT 216 a 418 c 390 g 242 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 3,56e-53 Length: 1266  
 Score: 1064.00 Matches: 184  
 Percent Similarity: 80.40% Conservative: 17  
 Best Local Similarity: 73.60% Mismatches: 49  
 Query Match: 73.89% Indels: 0  
 DB: 6 Gaps: 0

US-10-010-408-2 (1-250) x AX464186 (1-1266)

QY 1 MetArgGlySerProLeuIleHisLeuLeuAlaThrSerPheLeuCysLeuSerMet 20  
 Db 10 ATGAGAGGACACCGGAGAGACCCACCTCTGGCTTCTCCCTCTCTCTCTCTCTCTCTCTCAAG 69

QY 21 ValCysAlaGlnLeuCysArgThrProCysThrCysProThrThrProProGlnCysPro 40  
 Db 70 GTGGCTACCCAGCTGTGCCGACACCATGTACCTGCCCTGCCCTGCCCTGCCCTGCCCTGCCG 129

QY 41 GlnGlyValProLeuValLeuAspGlyCysGlyCysLysValCysValCysAlaArgLeu 60  
 Db 130 CTGGAGTACCCCTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGT 189

QY 61 GlyGluSerCysAspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnPro 80  
 Db 190 GGGAGGCCCTCGGACCAACTCCAGCTCTGCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCC 249

QY 81 GlyAlaGlyProGlyGlyHisGlyAlaValCysLeuLeuAspGluAspGlySerCys 100  
 Db 250 GGGCAGGACCCGCTGTGGCGGGGGGCCCTGTGCTCTGTGGCAGAGAGCAGCAGCAGCTGT 309

QY 101 GluValAsnGlyArgTyrLeuAspGlyGluThrPheLeuAspGlyGluThrProAsnCysArgValLeu 120  
 Db 310 GAGGTGACGGCCGCTGTATCGGAGGAGGAGACCTTCCAGCCCCCTGCTGCTGCTGCTGCTGCTG 369

QY 121 CysArgCysAspAspGlyGlyPheThrCysLeuProLeuCysSerGluAspValArgLeu 140  
 Db 370 TGGCGCTGCAGAGCGCGCGCTTACCTGCCTGCGCTGTGCAGCAGGATGTGCGGCTG 429

QY 141 ProSerTrpAspCysProArgProLysArgIleGlnValProGlyLysCysCysProGlu 160  
 Db 430 CCCAGCTGGGACTGCCCCACCCAGGAGGAGGTCTGAGGTCTGGGCAAGTGTGCTGCCCTGAG 489

QY 161 TrpValCysAspGlnGlyValThrProAlaIleGlnArgSerThrAlaGlnGlyHisGln 180  
 Db 490 TGGGTGTGGCGCAGAGGAGGGGACTGGGGACCCAGCCCTTCCAGCCCAAGAGCCCGAG 549

QY 181 LeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThrAla 200  
 Db 550 TTTTCTGGCCCTGTCT 609

QY 201 TrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGlnAsn 220  
 Db 610 TGGGAGCCCTGCTGCACACCTGTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 669

QY 221 ArgPheCysGlnLeuGluLeuGlnArgLeuCysLeuProArgProCysLeuAlaAla 240  
 Db 670 CGCTTCTGCCGAGTGGAGACCCAGCCGCGCTGTGCCTGTCCAGGCCCTGCCACCCCTCC 729

QY 241 ArgSerHisSerSerTrpAsnSerAlaPhe 250  
 Db 730 AGGGTCCGACGTCCACAAACAGTGCCTTC 759

RESULT 8  
 AF083500  
 LOCUS  
 DEFINITION Homo sapiens connective tissue growth factor-like protein  
 precursor, mRNA, complete cds.  
 ACCESSION AF083500  
 VERSION AF083500.1 GI:3462835  
 KEYWORDS  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
 AUTHORS Kumar,S., Hand,A.T., Connor,J.C., Dodds,R.A., Ryan,P.J.,  
 Trill,J.G., Fisher,S.M., Slemmon,J.R., Lipshutz,D.B.,  
 Bartholomew,V., James,I.E., Klemm,D.J., Gowen,M. and Lee,J.C.  
 TITL Identification and cloning of CTGF-L from human osteoblasts, a  
 novel cysteine rich protein containing an IGF binding domain  
 JOURNAL Bone 23 (5), S240 (1998)  
 REFERENCE 2 (bases 1 to 1283)  
 AUTHORS Kumar,S.  
 TITLE Direct Submission  
 JOURNAL Submitted (11-AUG-1998) Bone & Cartilage Biology, UW 2109,  
 SmithKline Beecham, 709 Swedeland Road, King of Prussia, PA 19406,  
 USA

FEATURES  
 source Location/Qualifiers  
 1. 1283  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="20"  
 /map="20q12-q13"  
 /celltype="primary osteoblast"  
 9. 761  
 /note="CTGF-L; encodes IGF binding (IGFB), Von Willebrand  
 factor type C (VWC) and thrombospondin type I (TSPI)  
 domains; member of the CCN (CTGF/cyrl6/nov) family; lacks  
 the fourth carboxy-terminal domain present in other  
 members of the CCN family"  
 /codon\_start=1  
 /product="connective tissue growth factor-like protein  
 precursor"  
 /protein\_id="AAC70350.1"  
 /db\_xref="GI:3462835"  
 /translation="MRGTPTKTHLAFSLLLKLLSKVTRTQCTPTCTCPWPPRCPGLGVP  
 ILVDG@CCRCVRCARLGEPCDQLHVCDSAGSLVCGPGGGRGALCLLAEDDSSCEV



US-10-010-408-2 (1-250) x AF074604 (1-1309)

|  |     |  |     |
|--|-----|--|-----|
| QY   | 1   | MetArgGlySerProLeuIleHisLeuLeuAlaThrSerPheLeuCysLeuSerMet    | 20  |
| DB   | 7   | ATGAGAGCACACGAGAACCCACCTCTGGCCCTTCTCCTCTGCTCTCTCTCAAG        | 66  |
| QY   | 21  | ValCysAlaGlnLeuCysArgThrProCysThrCysProThrPheProGlnCysPro    | 40  |
| DB   | 67  | GTGGCTACCCAGCTGTGCCGACACCAATGATCTCTGCTTGGCCACTCCCGATGCCG     | 126 |
| QY   | 41  | GlnGlyValProLeuValLeuAspGlyCysGlyCysLysValCysAlaArgArgLeu    | 60  |
| DB   | 127 | CTGGAGTACCCCTGGTGTGGATGGCTGTGGCTGTGGCGGTATGTGCACGGCGGTG      | 186 |
| QY   | 61  | GlyGluSerCysAspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnPro | 80  |
| DB   | 187 | GGGAGGCCCTCGGACCAACTCCAGCTGTGGCAGCCGACGAGCGGCTGTCTGCCAGGCC   | 246 |
| QY   | 81  | GlyAlaGlyProGlyGlyHisGlyAlaValCysLeuAspGluAspAspGlySerCys    | 100 |
| DB   | 247 | GGGCGAGACCCGGTGGCGGGGGCCCTGTGGCTTGGCAGGACGACAGCAGCTGT        | 306 |
| QY   | 101 | GluValAsnGlyArgArgTyrLeuAspGlyGluThrPheLysProAsnCysArgValLeu | 120 |
| DB   | 307 | GAGGTGAACGCCCTGTATCGGAGGGGAGACCTTCAGCCCTCAGCAGCATCCGC        | 366 |
| QY   | 121 | CysArgCysAspAspGlyGlyPheThrCysLeuProLeuCysSerGluAspValArgLeu | 140 |
| DB   | 367 | TGCCCTCGAGGACGGCGCTTCCCTGGCTGGCTGTGCAGCAGCATGTGGCGGTG        | 426 |
| QY   | 141 | ProSerTrpAspCysProArgProLysArgIleGlnValProGlyLysCysCysProGlu | 160 |
| DB   | 427 | CCAGCTGGGACTGCCCCACCCAGGAGGGTCTGAGGTCTGGGCAAGTGTCTGCCCTGAG   | 486 |
| QY   | 161 | TrpValCysAspGlnGlyValThrProAlaIleGlnArgSerThrAlaGlnGlyHisGln | 180 |
| DB   | 487 | TGGGTGTGGCCCAAGAGAGGGGACTGGGACCCAGCCCTTCCAGCCCAAGAGCCCGAG    | 546 |
| QY   | 181 | LeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThrAla | 200 |
| DB   | 547 | TTTTCTGGCCCTGTCTCTCTCCCTGCCCTGGTGTCCCTGCCACAGATGGAGCGGCC     | 606 |
| QY   | 201 | TrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrargValSerAsnGlnAsn | 220 |
| DB   | 607 | TGGGAGCCCTCTCGACCACTGTGGCTGGGATGGCCACCGGCTGTCCCAACAGAAC      | 666 |
| QY   | 221 | ArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProArgProCysLeuAlaAla | 240 |
| DB   | 667 | CGCTCTCCGACTGGAGACCCAGCCGCGCTGTGCTGTCCAGGCCCTGCCACCCCTCC     | 726 |
| QY   | 241 | ArgSerHisSerSerTrpAsnSerAlaPhe                               | 250 |
| DB   | 727 | AGGGGTGCGACTCCACAAACAGTGCCTTC                                | 756 |
| RESULT 12  |     |  |     |
| AF100780   |     |  |     |
| LOCUS  |     |  |     |
| DEFINITION   |     |  |     |
| Homo sapiens connective tissue growth factor related protein WISP-5    |     |  |     |
| (WISP-5) mRNA, complete cds.   |     |  |     |
| AF100780   |     |  |     |
| VERSION  |     |  |     |
| KEYWORDS   |     |  |     |
| SOURCE   |     |  |     |
| ORGANISM   |     |  |     |
| Homo sapiens.  |     |  |     |
| Homo sapiens   |     |  |     |
| Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;      |     |  |     |
| Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.            |     |  |     |
| 1 (bases 1 to 1427)  |     |  |     |
| PENNICK, D., SWANSON, T.A., WELSH, J.W., ROY, M.A., LAWRENCE, D.A.,    |     |  |     |
| LEE, J., BRUSH, J., TANEYHILL, L.A., DEUEL, B., LEW, M., WATANABE, C., |     |  |     |
| COHEN, R.L., WELLMAN, M.F., FINLEY, G.G., QUIRKE, P., GODDARD, A.D.,   |     |  |     |
| HILLAN, K.J., GURNEY, A.L., BOSTEIN, D. and LEVINE, A.J.               |     |  |     |
| WISP genes are members of the connective tissue growth factor          |     |  |     |
| family that are up-regulated in wnt-1-transformed cells and            |     |  |     |
| aberrantly expressed in human colon tumors                             |     |  |     |

JOURNAL PROC. Natl. Acad. Sci. U.S.A. 95 (25), 14717-14722 (1998)  
MEDLINE 99061933  
PUBMED 9843955  
REFERENCE 2 (bases 1 to 1427)  
AUTHORS Pennica,D.  
TITLE Direct Submission  
JOURNAL Submitted (23-OCT-1998) Molecular Oncology, Genentech, Inc., 1 DNA  
Way, South San Francisco, CA 94080, USA

FEATURES  
source Location/Qualifiers  
1..1427  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="20"  
/map="20q12-q13.1"  
/tissue\_type="lung"  
/dev\_stage="fetus"  
1..1427  
/gene="WISP2"  
156..908  
/gene="WISP2"  
/codon\_start=1  
/product="connective tissue growth factor related protein  
WISP-2"  
/protein\_id="AAC96322.1"  
/db\_xref="GI:4028583"  
/translocation="MRGTPKTHLLAFSLCLLSKVRLQLCPTCTCPWPPrCPLGVV  
LVLDGGCCRCVACRAGEPCDQLHVCDASQGLVCQPAGPGRGALCLLAEADSSCEV  
NGLYREGFTFPHCSIRCRCEGDGTFCVPLCEDSEVRLPMSDCPHPRRVVLCKCQPE  
WVGCGGGGLGTOPLPAGQGQFSGLVSLPPGVPCPEWSMTANGPCSTTCGLGMATRYNS  
QNFRCLERTORRLCLRCPDRPSGRSPONSAP"

BASE COUNT 261 a 468 c 417 g 280 t 1 others  
ORIGIN

Alignment Scores:  
Pred. No.: 4,01e-53 Length: 1427  
Score: 1064.00 Matches: 184  
Percent Similarity: 80.40% Conservative: 17  
Best Local Similarity: 73.60% Mismatches: 49  
Query Match: 73.89% Indels: 0  
DB: Gaps: 0

US-10-010-408-2 (1-250) x AF100780 (1-1427)

|    |     |  |     |
|----|-----|--|-----|
| Qy | 1   | MetArgGlySerProLeuIleHisLeuLeuAlaThrSerPheLeuCysLeuLeuSerMet   | 20  |
| Db | 156 | ATGAGAGGCACACCGAAGACCACCTCTGGCCTTCTCCCTCTCTCCCTCTCTCAAG        | 215 |
| Qy | 21  | ValCysAlaGlnLeuCysArgThrProCysThrCysProThrphrProProGlnCysPro   | 40  |
| Db | 216 | GTCGCTACCCAGCTGTGCCCGCACCATGATACCTGCCCTGGCCACCTCCCGATGCCG      | 275 |
| Qy | 41  | GlnGlyValProLeuValLeuAspGlyCysGlyCysCysValCysAlaArgArgLeu      | 60  |
| Db | 276 | CTGGAGTAGCCCTGGTGCTGGATGGCTGTGGCTGTGGCTGGCTGGTGGTGGCGGCTG      | 335 |
| Qy | 61  | GlyGluSerCysAspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnPro   | 80  |
| Db | 336 | GGGAGGCCCTGCGACCAACTCCAGCTGCGACGCCAGCCAGGGCCCTGGTCTGCCAGCCC    | 395 |
| Qy | 81  | GlyAlaGlyProGlyGlyHisGlyAlaValCysLeuLeuAspGluAspGlySerCys      | 100 |
| Db | 396 | GGGCGAGGACCCGCTGGCGGGGGCCCTGTGCCCTCTGGCAGAGGACGACGAGCTGT       | 455 |
| Qy | 101 | GluValAsnGlyArgArgTyrrLeuAspGlyGluThrPhelysProAsnCysArgValLeu  | 120 |
| Db | 456 | GAGGTGAACGGCCGCTTTATTCGCGAGGGAGACCTTCACGCCCTCTCCAGCATCCGC      | 515 |
| Qy | 121 | CysArgCysAspGlyGlyPheThrCysLeuProLeuCysSerGluAspValArgLeu      | 140 |
| Db | 516 | TGCCGCTGCGAGGACGGCGGCTTCACCTGCTGCCGCTGTGCAGGAGATGTCGGCTG       | 575 |
| Qy | 141 | ProSerTrpAspCysProArgProLyisArgIleGlnValProGlyLyisCysCysProGlu | 160 |

```

Db      576  CCAGCTGGGACTGCCCCACCCAGGAGGGTCGAGGTCTCTGGGCAAGTGTGCGCTGAG 635
QY      161  TrpValCysAspGlnGlyValThrProAlaIleGlnArgSerThrAlaGlnGlyHisGln 180
Db      636  TGGGTGTGGCCAGGAGGGGACTTGGGAGCCACCCCTTCCAGCCCAAGACCCAG 695
QY      181  LeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThrAla 200
Db      696  TTTTCTGGCCTTGTCTCTCTCCCTGGCCCTGGTGGTCTCTGGCCAGAAATGGAGCAGCGCC 755
QY      201  TrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGlnAsn 220
Db      756  TGGGAGCCCTGCTCAGCACCTGTGGCTGGGATGGCCACCCGGGTGTCCAAACAGAAC 815
QY      221  ArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProArgProCysLeuAlaAla 240
Db      816  CGCTTCTGGGACTGGAGACCCAGCGCCGCTGTGCTGCTCCAGGCGCTTCCACCCCTCC 875

RESULT 13
BC017782
LOCUS   BC017782
DEFINITION Homo sapiens, WNT1 inducible signaling pathway protein 2, clone
ACCESSION BC017782
VERSION   BC017782.1
KEYWORDS  MGC.
SOURCE    Homo sapiens.
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1450)
Strausberg, R.
Direct Submission
Submitted (03-DEC-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapps-remail.nih.gov
Tissue Procurement: CLONTECH
CDNA Library Preparation: CLONTECH Laboratories, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LNLNL)
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www.shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNLNL at: http://image.llnl.gov
Series: IRAL Plate: 36 Row: m Column: 3
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 4507922.
FEATURES
Location/Qualifiers
1..1450
/organism="Homo sapiens"
/db_xref="LocusID:8839"
/db_xref="taxon:9606"
/clone="MGC:22271 IMAGE:4691574"
/tissue_type="Lung"
/clone_lib="NIH_MGC_77"
/lab_host="DH10B"
/note="Vector: pDNR-LIB"
165..917
/codon_start=1
/product="WNT1 inducible signaling pathway protein 2"
/protein_id="AAH17782.1"

CDS

```

```

/db_xref="GI:17389483"
/translation="MRGTPKTHLLAFSLLLSKVRLQLCPTPTCTCPMPPPRCPLGVP
LVLDGCGCCRCVARRLGPQDOLHVCDSQGLVCPGAGPGRRGALCLLAEDESCEV
NGRLYREGTETQPHCSIRCRDGGCTCTVPLCSQEDVRLPSWDCPHRRVEVLGKCCPE
WVCGGGSLGTQPLPAQGPQPSGLYSSLPPLGVPCEPSTANGPCSTTCGLGNATRVSN
QNRFGLETORRLCLSRFCPPSPRSQNSAF"
BASE COUNT 272 a 457 c 441 g 280 t
ORIGIN
Alignment Scores:
Pred. No.: 4,08e-53 Length: 1450
Score: 1064.00 Matches: 184
Percent Similarity: 80.40% Conservative: 17
Best Local Similarity: 73.60% Mismatches: 49
Query Match: 73.89% Indels: 0
DB: 9 Gaps: 0
US-10-010-408-2 (1-250) x BC017782 (1-1450)
QY 1 MetArgGlySerProLeuIleHisLeuAlaThrSerPheLeuCysLeuLeuSerMet 20
Db 165 ATGAGAGGACACGACGAGACCCACCTCTGGGCTTCTCCCTCTCTCTCTCAAG 224
QY 21 ValCysAlaGlnLeuCysArgThrProCysThrCysProThrThrProGlnCysPro 40
Db 225 GTGCGTACCCAGCTGTGCCGACACCATGTACTGCCCTGGCCACCTCCCGATGCCG 284
QY 41 GlnGlyValProLeuValLeuAspGlyCysGlyCysCysValCysValCysAlaArgLeu 60
Db 285 CTGGGAGTACCCCTGTGTGGATGGCTGTGGCTGTCTGCCGGTATGTGCAGCGGGCTG 344
QY 61 GlyGluSerCysAspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnPro 80
Db 345 GGGGAGCCCTGGACCACTCCACGCTCTGCGAGCGCCAGGAGGCTGTGCTGCGAGCCC 404
QY 81 GlyAlaGlyProGlyGlyHisGlyAlaValCysLeuLeuAspGluAspAspGlySerCys 100
Db 405 GGGGAGAGACCCGGTGGCGGGGCGCTGTGCTCTTGGCAGAGAGGACGACAGCTGT 464
QY 101 GluValAsnGlyArgArgTyrLeuAspGlyGluThrPheLeuProAsnCysArgValLeu 120
Db 465 GAGGTGAACGGCGCGCTGTATCGGAAGGGAGAGCTTCCAGCCCATGACGATCGCG 524
QY 121 CysArgCysAspGlyGlyPheThrCysLeuProLeuCysSerGluAspValArgLeu 140
Db 525 TCCCGCTGGAGGACGGCGGTTCACCTGCGTGGCGCTGTGCGAGGAGATGTGCGGCTG 584
QY 141 ProSerTrpAspCysProArgProlyArgIleGlnValProGlyLysCysProGlu 160
Db 595 CCCAGCTGGGACTGCCCCACCCAGAGGGTGTGAGGTCTTGGGCAAGTGTGCTGCTGAG 644
QY 161 TrpValCysAspGlnGlyValThrProAlaIleGlnArgSerThrAlaGlnGlyHisGln 180
Db 645 TGGGTGTGGGAGGAGGGGACTGGGGAGCCACCCCTTCCAGCCCAAGACCCAG 704
QY 181 LeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThrAla 200
Db 705 TTTTCTGGCCTTGTCTCTCTCCCTGGCCCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 764
QY 201 TrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGlnAsn 220
Db 765 TGGGAGCCCTGTGTCACCACTGTGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 824
QY 221 ArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProArgProCysLeuAlaAla 240
Db 825 CGCTTCTGGGACTGGAGACCCAGCGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 884
QY 241 ArgSerHisSerSerTrpAsnSerAlaPhe 250
Db 885 AGGGGTGCGAGTCCACAAAACAGTGCCTTC 914
RESULT 14
AR210337

```



```

Db      492 TGGGTGTGGGCCAAGGAGGGGGACTGGGGACCAAGCCCTT-----CCAGCCCAAGG 542
QY      178 yHisGlnLeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSe 198
Db      543 ACCCAGTTTCTGGCCTTGTCTCTCCCTGCCCTGGTGTCCCTGCCCAAGATGGAG 602
QY      198 rThrAlaTrpGlyProCysSerThrThrCysGlyLeuGlyLeuAlaThrArgValSerAs 218
Db      603 CACGGCCTGGGACCCCTGCTCGACCACTGTGGGTGGCATGGCCACCCGGGTGTCCAA 662
QY      218 nGlnAsnArgPheCysGlnLeuGluIleGlnArgLeuCysLeuProArgProCysLe 238
Db      663 CCAGAACCGCTTCTGCCGACTGGAGACCCAGCGCGCTGTGCTGTCCAGGCCCTGCC 722
QY      238 uAlaAlaArgSerHisSerSerTrpAsnSerAlaPhe 250
Db      723 ACCCTCCAGGGGTCCGAGTCCACAAACAGTGCCTTC 759

```

Search completed: August 1, 2003, 04:21:56  
Job time : 2649 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 1, 2003, 01:17:19 ; Search time 237 Seconds  
(without alignments)  
2375.526 Million cell updates/sec

Title: US-10-010-408-2

Perfect score: 1440

Sequence: 1 MRGSLPHLLATSFLLLSM.....LCLPRPCLAAARSHSSWNSAF 250

Scoring table:

BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh  
-O=/cgn2.1/USPRO.spool/US10010408/runat\_25072003\_101855\_17876/app\_query.fasta\_1.391  
-DB=N\_Geneseq\_101002\_QFNT-fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blom62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFT=ptc -NORM=exp -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10010408 -CGEN\_1\_1\_263 @runat\_25072003\_101855\_17876 -ICPU=6 -ICPU=3  
-NO\_MMAP -LARGOQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=170 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N\_Geneseq\_101002:\*  
1: /SID2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:\*  
2: /SID2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:\*  
3: /SID2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:\*  
4: /SID2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:\*  
5: /SID2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:\*  
6: /SID2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:\*  
7: /SID2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:\*  
8: /SID2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:\*  
9: /SID2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:\*  
10: /SID2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:\*  
11: /SID2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:\*  
12: /SID2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:\*  
13: /SID2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:\*  
14: /SID2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:\*  
15: /SID2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:\*  
16: /SID2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:\*  
17: /SID2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:\*  
18: /SID2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:\*  
19: /SID2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:\*  
20: /SID2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:\*  
21: /SID2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:\*  
22: /SID2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:\*  
23: /SID2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:\*  
24: /SID2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score  | Query Match % | Length | DB ID | Description        |
|------------|--------|---------------|--------|-------|--------------------|
| 1          | 1440   | 100.0         | 753    | 20    | Rat HICP polypepti |
| 2          | 1440   | 100.0         | 1708   | 20    | AAZ07517           |
| 3          | 1323   | 91.9          | 681    | 20    | AAZ07516           |
| c          | 1308.5 | 90.9          | 753    | 20    | Rat HICP mature po |
| 5          | 1308.5 | 90.9          | 1734   | 20    | Mouse WISP-2 prote |
| c          | 1064   | 73.9          | 750    | 20    | Mouse WISP-2 prote |
| 7          | 1064   | 73.9          | 1257   | 20    | Human WISP-2 prote |
| 8          | 1064   | 73.9          | 1266   | 21    | EGF-like homologue |
| 9          | 1064   | 73.9          | 1266   | 22    | Human CDNA sequenc |
| 10         | 1064   | 73.9          | 1266   | 22    | Human CDNA sequenc |
| 11         | 1064   | 73.9          | 1266   | 22    | PRO261 coding sequ |
| 12         | 1064   | 73.9          | 1285   | 19    | Human angiogenesis |
| 13         | 1064   | 73.9          | 1293   | 20    | Human connective t |
| 14         | 1064   | 73.9          | 1309   | 22    | Human WISP-2 prote |
| 15         | 1064   | 73.9          | 1337   | 22    | Connective tissue  |
| 16         | 1064   | 73.9          | 1352   | 22    | Human secreted pro |
| 17         | 1064   | 73.9          | 1522   | 20    | Human secreted pro |
| 18         | 1051   | 73.0          | 1267   | 21    | Human growth facto |
| 19         | 1047   | 72.7          | 738    | 20    | Human PRO261 cDNA  |
| 20         | 1045.5 | 72.6          | 841    | 20    | Human WISP-2 prote |
| 21         | 564    | 39.2          | 2075   | 16    | Human WISP-2 prote |
| 22         | 564    | 39.2          | 2075   | 18    | Connective tissue  |
| 23         | 564    | 39.2          | 2075   | 18    | Connective tissue  |
| 24         | 564    | 39.2          | 2075   | 19    | Connective tissue  |
| 25         | 564    | 39.2          | 2075   | 20    | Human connective t |
| 26         | 564    | 39.2          | 2075   | 21    | Human connective t |
| 27         | 564    | 39.2          | 2075   | 21    | Human connective t |
| 28         | 564    | 39.2          | 2075   | 21    | Human connective t |
| 29         | 564    | 39.2          | 2075   | 22    | Human connective t |
| 30         | 564    | 39.2          | 2075   | 22    | Human connective t |
| 31         | 564    | 39.2          | 2075   | 22    | Human benign prost |
| 32         | 564    | 39.2          | 2312   | 22    | ABK64553           |
| 33         | 564    | 39.2          | 2312   | 22    | Nucleotide sequenc |
| 34         | 550.5  | 38.2          | 2267   | 22    | Human shear stress |
| 35         | 550.5  | 38.2          | 2267   | 22    | AAH02914           |
| 36         | 550.5  | 38.2          | 2267   | 22    | Murine Fisp12 cDNA |
| 37         | 550.5  | 38.2          | 2330   | 13    | Mouse fibroblast s |
| 38         | 548.5  | 38.1          | 2338   | 20    | Gene for beta-IG-M |
| 39         | 542.5  | 37.7          | 2350   | 21    | Rat sequence diffe |
| 40         | 541.5  | 37.6          | 1975   | 14    | Rat connective tis |
| 41         | 532    | 36.9          | 4231   | 24    | Clone 2-4-7 encodi |
| 42         | 520.5  | 36.1          | 684    | 14    | AAQ36031           |
| 43         | 511    | 35.5          | 4212   | 20    | Chicken nov coding |
| 44         | 511    | 35.5          | 4214   | 18    | Human DNA sequenc  |
| 45         | 511    | 35.5          | 4214   | 18    | Chicken nov gene f |
|            |        |               |        |       | Connective tissue  |
|            |        |               |        |       | Human connective t |
|            |        |               |        |       | AAAT59618          |
|            |        |               |        |       | AAAT58534          |

ALIGNMENTS

RESULT 1  
AAZ07517

ID AAZ07517 standard; cDNA: 753 BP.

AC AAZ07517;

DT 26-NOV-1999 (first entry)

DE Rat HICP polypeptide coding sequence.

DE Heparin-induced CCN-like protein; HICP; cell-associated activity; ss;

DE cardiovascular disorder; aberrant cell proliferation; fibrotic disorder.

OS Rattus sp.

PN WO9947556-A2.

XX 23-SEP-1999.

PD

XX 18-MAR-1999; 99WO-US05999.  
 XX 19-MAR-1998; 98US-0044273.  
 XX (TUFT ) TUFTS COLLEGE.  
 XX Castellot JJ;  
 XX WPI: 1999-562060/47.  
 XX P-PSDB; AAY27434.  
 XX Nucleic acid sequences encoding rat heparin-induced CCN-like protein,  
 PT used in methods to identify modulators or in diagnostic applications  
 PT  
 XX Claim 5; Fig 1; 108pp; English.  
 XX The invention provides a rat heparin-induced CCN-like protein (HICP)  
 CC protein. Agents that stimulate or inhibit HICP protein activity or  
 CC expression, antisense HICP nucleic acid molecules and HICP antibodies,  
 CC can be used to modulate cell-associated activity. HICP modulators can be  
 CC used to treat disorders characterized by aberrant HICP protein activity  
 CC or expression. Probes capable of hybridizing to HICP mRNA or antibodies  
 CC specific for HICP can be used to detect HICP activity in a biological  
 CC sample. HICP can be used to treat disorders, such as a cardiovascular or  
 CC fibrotic disorder, characterized by aberrant cell proliferation. The  
 CC present sequence represents the coding sequence of rat HICP.  
 XX  
 XX Sequence 753 BP; 132 A; 235 C; 230 G; 156 T; 0 other;  
 SQ

Alignment Scores:  
 Pred. No.: 1.07e-77 Length: 753  
 Score: 1440.00 Matches: 250  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 20 Gaps: 0

US-10-010-408-2 (1-250) x AAZ07517 (1-753)

QY 1 MetArgGlySerProLeuIleHisLeuLeuAlaThrSerPheLeuCysLeuSerMet 20  
 DB 1 ATGAGGGGAGCCCACTGATCCATCTCTGGCCACTTCTCTCTGCTTCTCTCAATG 60  
 QY 21 ValCysAlaGlnLeuCysArgThrProCysThrCysProThrThrProGlnCysPro 40  
 DB 61 GTGTGTGCCAGCTGTGCGGACACCTGTACTCTCTCTGGACACCCAGTGCCCA 120  
 QY 41 GlnGlyValProLeuValLeuAspGlyCysGlyCysLysValCysAlaArgLeu 60  
 DB 121 CAGGGGGTACCCCTGGTGGATGGCTGTGCTGCTGTAAAGTGTGTGACGGAGGTG 180  
 QY 61 GlyGluSerCysAspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnPro 80  
 DB 181 GGGAGTCTGTGGACCACTGCTGATGCTGTGGACCCAGCCAGGGCTGTGTGTGAGCCT 240  
 QY 81 GlyAlaGlyProGlyGlyHisGlyAlaValCysLeuLeuAspGluAspGlySerCys 100  
 DB 241 GGGGAGCCCTGTGGCGCCAGTGGGCTGTGTCTCTGGATGAGATGACGGTAGCTGT 300  
 QY 101 GluValAsnGlyArgArgTyrLeuAspGlyGluThrPheLysProAsnCysArgValLeu 120  
 DB 301 GAGGTGAATGGCGCGAGGTACCTGTGATGGAGAGACCTTTAAACCAATTGCAGGGTCTG 360  
 QY 121 CysArgCysAspAspGlyGlyPheThrCysLeuProLeuProCysSerGluAspValArgLeu 140  
 DB 361 TCCCTCTGTATGACGGTGGCTTACCTGCTCTGCGCTGTGTGAGTGAGGATGTGGGCTG 420  
 QY 141 ProSerTrpAspCysProArgProLysArgIleGlnValProGlyLysCysProGlu 160  
 DB 421 CCCAGCTGGAGTCCCGCCAGCCCAAGAAATACAGGTGCCAGGAAGTGTGCCCCAG 480  
 QY 161 TrpValCysAspGlnGlyValThrProAlaIleGlnArgSerThrAlaGlnGlyHisGln 180

DB 481 TGGGTATGTGACCGAGGAGTGACACCGCGATCCACGCGCTCCACGCGCAAGGACCA 540  
 QY 181 LeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThrAla 200  
 DB 541 CTTTCTGCCCTTGTCTACCTCCCTGCTGTATGCTCTCTGTCCAATTTGGAGCACGCC 600  
 QY 201 TrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGlnAsn 220  
 DB 601 TGGGGCCCTGCTCAACCACTGTGGCTGGCATAGCCACCGAGTGTCCAACCCAGAAC 660  
 QY 221 ArgPheCysGlnLeuGluIleGlnArgLeuCysLeuProArgProCysLeuAlaAla 240  
 DB 661 CGATTCTGCAAGTGGAGATCCACGCGCGCTGTGTCTGCCAGACCCCTGCCCTGGCAGCC 720  
 QY 241 ArgSerHisSerSerTrpAsnSerAlaPhe 250  
 DB 721 AGGAGCCACAGCTCATGGAACAGTCTTTC 750

RESULT 2  
 AAZ07516  
 ID AAZ07516 standard; cDNA; 1708 BP.  
 XX AAZ07516;  
 AC AAZ07516;  
 XX 26-NOV-1999 (first entry)  
 DT  
 XX Rat HICP polypeptide encoding cDNA.  
 DE  
 XX Heparin-induced CCN-like protein; HICP; cell-associated activity; as;  
 KW cardiovascular disorder; aberrant cell proliferation; fibrotic disorder.  
 KW  
 XX Rattus sp.  
 OS  
 XX WO9947556-A2.  
 PN  
 XX 23-SEP-1999.  
 PD  
 XX 18-MAR-1999; 99WO-US05999.  
 PF  
 XX 19-MAR-1998; 98US-0044273.  
 PR  
 XX (TUFT ) TUFTS COLLEGE.  
 PA  
 XX Castellot JJ;  
 PI  
 XX WPI: 1999-562060/47.  
 DR  
 XX P-PSDB; AAY27434.  
 DR  
 XX Nucleic acid sequences encoding rat heparin-induced CCN-like protein,  
 PT used in methods to identify modulators or in diagnostic applications  
 PT  
 XX Claim 2; Fig 1; 108pp; English.  
 PS  
 XX This cDNA encodes a rat heparin-induced CCN-like protein (HICP) protein.  
 CC Agents that stimulate or inhibit HICP protein activity or expression,  
 CC antisense HICP nucleic acid molecules and HICP antibodies, can be used to  
 CC modulate cell-associated activity. HICP modulators can be used to treat  
 CC disorders characterized by aberrant HICP protein activity or expression.  
 CC Probes capable of hybridizing to HICP mRNA or antibodies specific for  
 CC HICP can be used to detect HICP activity in a biological sample. HICP  
 CC can be used to treat disorders, such as a cardiovascular or fibrotic  
 CC disorder, characterized by aberrant cell proliferation.  
 XX  
 XX Sequence 1708 BP; 362 A; 486 C; 478 G; 382 T; 0 other;  
 SQ

Alignment Scores:  
 Pred. No.: 2.54e-77 Length: 1708  
 Score: 1440.00 Matches: 250  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 20 Gaps: 0





[illegible]



XX This sequence encodes the EGF-like homologue PRO261.  
CC The invention relates to antibodies (Ab) that bind to any of the  
CC polypeptides (I) designated PRO187; PRO533; PRO214; PRO240; PRO211;  
CC PRO230; PRO261; PRO246 or EBAF-2. The Ab, or other agents that inhibit  
CC expression and/or activity of (I) are used: (i) to inhibit growth of  
CC tumours; and (ii) as diagnostic/prognostic reagents for detection or  
CC quantification of (I) in cells or tissues, by standard immunoassays, with  
CC overexpression being indicative of cancer. For therapeutic use, the Ab  
CC may be conjugated to a toxin, chemotherapeutic agent or radioisotope.  
CC Genes expressing (I), many of which are growth factor homologues, are  
CC overexpressed in some cases of cancer.

XX Sequence 1257 BP; 215 A; 416 C; 385 G; 241 T; 0 other;

Alignment Scores:  
Pred. No.: 4.86e-55 Length: 1257  
Score: 1064.00 Matches: 184  
Percent Similarity: 80.40% Conservative: 17  
Best Local Similarity: 73.60% Mismatches: 49  
Query Match: 73.89% Indels: 0  
DB: 20 Gaps: 0

US-10-010-408-2 (1-250) x AAX28435 (1-1257)

Qy 1 MetArgGlySerProLeuIleHisLeuLeuAlaThrSerPheLeuLeuSerMet 20  
Db 1 ATGAGAGCACACCGAAGACCCACCTCTGCGCCTCTCCCTCTCTCTCAAG 60  
Qy 21 ValCysAlaGlnLeuCysArgThrProCysThrCysProThrProGlnCysPro 40  
Db 61 GTGGGTACCCACCTGTGCGCCACACCATGTACCTGCGCCCTGCCCGATGCCG 120  
Qy 41 GlnGlyValProLeuValLeuAspGlyCysGlyCysValCysAlaArgLeu 60  
Db 121 CTGGGATACCCCTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGT 180  
Qy 61 GlyGluSerCysAspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnPro 80  
Db 181 GGGAGCCCTGCGACCACTCCACGTCTGCGACGCCAGCGGCGCTGTGTCGACGCC 240  
Qy 81 GlyAlaGlyProGlyGlyHisGlyAlaValCysLeuLeuAspGluAspGlySerCys 100  
Db 241 GGGGAGACCCGCTGGCGGGGCGCTGTGCTGTGGCAGAGAGAGAGAGAGAGAGTGT 300  
Qy 101 GluValAsnGlyArgTyrLeuAspGlyGluThrPheLeuPheLeuPheLeuValLeu 120  
Db 301 GAGGTGAACGCCCTGTATCGGGAAGGGGAGAGACCTTCAGCCCCCTGCGAGCATCCGC 360  
Qy 121 CysArgCysAspGlyGlyPheThrCysLeuProLeuCysSerGluAspValArgLeu 140  
Db 361 TGCCGCTGGAGGAGCGGCTTCCCTGCTGCGCTGTGTCAGGAGGATGTGGGCTG 420  
Qy 141 ProSerTrpAspCysProArgProlyArgIleGlnValProGlyLysCysCysProGlu 160  
Db 421 CCCAGCTGGGACTGCCCCACCCAGGAGGCTGAGTCTCTGGGCAAGTGTGCCCTGAG 480  
Qy 161 TrpValCysAspGlnGlyValThrProAlaIleGlnArgSerThrAlaGlnGlyHisGln 180  
Db 481 TGGGTGTGCGGCCAAGGAGGGGGAGTGGGGACCCAGCCCTTCAGCCCAAGGACCCAG 540  
Qy 181 LeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThrAla 200  
Db 541 TTTTCTGCCTGTCTCTCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600  
Qy 201 TrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGlnAsn 220  
Db 601 TGGGACCTCTCTGACACACCTGTGGGTGGGCATGGCCACCCGCGGTGTCCAACAGAAC 660  
Qy 221 ArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProArgProCysLeuAlaAla 240  
Db 661 CGCTTCTGCCAGCTGGAGACCCAGCGCGCTGTGCTGTCCAGCGCCCTGCCACCTCC 720

Qy 241 ArgSerHisSerSerTrpAsnSerAlaPhe 250  
Db 721 AGGGCTCGCAGTCCACAAACAGTCCTTC 750  
RESULT 8  
AAA30048  
ID AAA30048 standard; cDNA; 1266 BP.  
XX  
AC AAA30048;  
XX  
DT 09-AUG-2000 (first entry)  
XX  
DE Human PRO261 nucleotide sequence.  
XX  
KW Antibody; PRO187; PRO533; PRO214; PRO240; PRO211; PRO230; PRO261; PRO246;  
KW PRO317; tumour growth inhibitor; cancer; diagnosis; treatment; human;  
KW cell growth; proliferation; growth factor; ADEPT;  
XX antibody dependent enzyme mediated prodrug therapy; ss.  
OS Homo sapiens.  
XX  
PN WO200015666-A2.  
XX  
PD 23-MAR-2000.  
XX  
PF 08-SEP-1999; 99WO-US20594.  
XX  
PR 10-SEP-1998; 98US-0099803.  
PR 10-SEP-1998; 98WO-US18824.  
XX  
PA (GETH ) GENENTECH INC.  
XX  
PI Goddard A, Gurney AL, Hillan KJ, Roy MA, Wood WI, Botstein D;  
XX WPI: 2000-271386/23.  
DR P-PSDB; AAY88573.  
XX  
New isolated antibodies which bind to specific polypeptides used for  
diagnosis and treatment of neoplastic cell growth and proliferation -  
Example 7; Fig 13; 200pp; English.

This sequence represents a human PRO261 nucleotide sequence. PRO261 is a growth factor. The invention relates to isolated antibodies which bind to a polypeptide. The "PRO" polypeptides are encoded by genes which are over expressed in the genome of tumour cells. Vectors and host cells comprising the nucleic acid encoding the antibodies are used in the production of the antibodies. The antibodies and nucleic acids encoding them are used for diagnosing a tumour in a mammal. The antibodies are used for inhibiting the growth of tumour cells and identifying compounds that inhibit a biological or immunological activity of and/or expression of a PRO187, PRO533, PRO214, PRO240, PRO211, PRO230, PRO261, PRO246 or PRO317 polypeptide. The antibody can be used in antibody dependent enzyme mediated prodrug therapy (ADEPT) by conjugating the antibody to a prodrug-activating enzyme which converts a prodrug to an anti-cancer drug. The antibodies can be fluorescently labelled and monitored by light microscopy, flow cytometry or fluorimetry for diagnosis and prognosis of tumours.

XX Sequence 1266 BP; 216 A; 418 C; 390 G; 242 T; 0 other;

Alignment Scores:  
Pred. No.: 4.9e-55 Length: 1266  
Score: 1064.00 Matches: 184  
Percent Similarity: 80.40% Conservative: 17  
Best Local Similarity: 73.60% Mismatches: 49  
Query Match: 73.89% Indels: 0  
DB: 21 Gaps: 0

US-10-010-408-2 (1-250) x AAA30048 (1-1266)

Qy 1 MetArgGlySerProLeuIleHisLeuLeuAlaThrSerPheLeuLeuSerMet 20  
Db 1 ATGAGAGCACACCGAAGACCCACCTCTGCGCCTCTCCCTCTCTCTCAAG 60







```
Db 370 TCCGCTCGAGGAGGGCGGCTTACCTGCGCTGCGCTGTCACGCGAGGATGTGCGGCTG 429
Qy 141 ProSerTirAspCysProArgProLysArgIleGlnValProGlyLysCysCysProGlu 160
Db 430 CCCAGCTGGAGTCCGCCACCCAGGAGGCTCGAGGCTCTGGGCAAGTGTGCTCCCTGAG 489
Qy 161 TrpValCysAspGlnGlyValThrProAlaIleGlnArgSerThrAlaGlnGlyHisGln 180
Db 490 TGGTGTGCGGCCAGGAGGAGGAGGCTGGGGAGCCAGCCCTTCCAGGCCAAGGACCCCG 549
Qy 181 LeuSerAlaLeuValThrProAlaSerAlaAlaProCysProAsnTirSerThrAla 200
Db 550 TTTTCTGCGCTGTCTCTCCCTGCCCTCGTGTGCTGCTGCCAGATGAGACGCGCC 609
Qy 201 TrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGlnAsn 220
Db 610 TGGGACCCCTGCTCGACACCTGTGGCTGGGCTGGCATGGCCACCCGGGTGTCCAACAGAAC 669
Qy 221 ArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProArgProCysLeuAlaAla 240
Db 670 CGCTTCTGCGGACTGGAGACCCAGCGCGCTGTGCTGTCTCCAGGCGCTGCCACCTCC 729
Qy 241 ArgSerHisSerTirAsnSerAlaPhe 250
Db 730 AGGGTCCGACTCCACAAAACAGTGCCTTC 759

RESULT 11
AAC97451
ID AAC97451 standard; cDNA; 1266 BP.
AC AAC97451;
DT
DE 28-FEB-2001 (first entry)
DE Human angiogenesis-associated protein PRO261 cDNA, SEQ ID NO:107.
KW Human; angiogenesis-associated protein; PRO; endothelial cell growth;
KW cardiac hypertrophy; cardiovascular disorder; endothelial disorder;
KW angiogenic disorder; atherosclerosis; osteoporosis; hypertension;
KW myocardial infarction; diabetic retinopathy; rheumatoid arthritis;
KW Crohn's disease; psoriasis; endometriosis; ulcer; wound healing; cancer;
KW Alzheimer's disease; Huntington's disease; stroke; drug screening;
KW gene therapy; transgenic animal; ss.
XX
OS Homo sapiens.
XX
PN WO200053753-A2.
XX
PD 14-SEP-2000.
XX
PF 05-JAN-2000; 2000WO-US00219.
XX
PR 08-MAR-1999; 99WO-US05028.
PR 12-MAR-1999; 99US-0123957.
PR 14-MAY-1999; 99US-0134287.
PR 02-JUN-1999; 99WO-US12252.
PR 23-JUN-1999; 99US-0141037.
PR 20-JUL-1999; 99US-0144758.
PR 26-JUL-1999; 99US-0145698.
PR 01-SEP-1999; 99WO-US20111.
PR 08-SEP-1999; 99WO-US20594.
PR 15-SEP-1999; 99WO-US21090.
PR 15-SEP-1999; 99WO-US21547.
PR 05-OCT-1999; 99WO-US23089.
PR 30-NOV-1999; 99WO-US28313.
PR 30-NOV-1999; 99WO-US28409.
PR 02-DEC-1999; 99WO-US28564.
PR 02-DEC-1999; 99WO-US28565.
XX
PA (GETH ) GENENTECH INC.
XX
PI Ashkenazi AJ, Baker KP, Ferrara N, Gerber H, Goddard A;
PI Godowski PJ, Gurney AL, Hillan KJ, Kuo SS, Mark MR, Marsters SA;
```

```
PI Paoni NF, Pitti RM, Watanabe CK, Williams PM, Wood WT;
XX
XX WPI: 2001-090793/10.
DR P-PSDB; AAB53084.
XX
XX New isolated nucleic acid for producing a PRO polypeptide, analyzing
PT genetic disorders and treating cardiovascular, endothelial or
PT angiogenic disorders, such as atherosclerosis, wounds or cancer -
PS Claim 58; Fig 41; 293pp; English.
XX
XX The invention relates to novel human angiogenesis-associated proteins
CC designated PRO proteins (AAB53064-B53097), and to nucleic acids encoding
CC PRO proteins. The invention also relates to vectors and host cells
CC comprising a PRO nucleic acid, the recombinant production of a PRO
CC protein, PRO antibodies specific for a PRO protein, fusion proteins
CC comprising a PRO protein, agonists or antagonists of a PRO protein, and
CC compounds which inhibit the expression of a PRO gene. The invention
CC additionally encompasses methods of identifying modulators of PRO
CC expression or activity; diagnosing a cardiovascular, endothelial or
CC angiogenic disorder, or a susceptibility to such a disorder by detecting
CC mutations in a PRO gene, or the expression level of a PRO gene within a
CC particular tissue; treating a cardiovascular, endothelial or angiogenic
CC disorder via the administration of a PRO protein, PRO nucleic acid, or
CC PRO agonist or antagonist; a retroviral gene therapy vector comprising a
CC PRO nucleic acid; and methods of inhibiting or stimulating endothelial
CC cell growth, cardiac hypertrophy or PRO-induced angiogenesis via the
CC administration of a PRO protein, or an agonist or antagonist thereof.
CC PRO nucleic acids, PRO proteins, antibodies against PRO proteins, PRO
CC agonists and PRO antagonists may be used as therapeutic agents to treat
CC cardiovascular, endothelial or angiogenic disorders, such as
CC atherosclerosis, osteoporosis, myocardial infarction, hypertension,
CC diabetic retinopathy, rheumatoid arthritis, Crohn's disease, Huntington's
CC endometriosis, ulcers, wounds, cancer, Alzheimer's disease, Huntington's
CC disease, or stroke. PRO nucleic acids are additionally useful in the
CC recombinant production of PRO proteins, as hybridisation probes to
CC screen libraries to isolate cDNAs with sequence identity to PRO proteins,
CC to map genes encoding PRO proteins, to analyse genetic disorders, and in
CC gene therapy. PRO nucleic acids can also be used to produce transgenic
CC animals useful for the development and screening of potential
CC therapeutic agents. The present sequence represents a cDNA encoding a PRO
CC protein of the invention.
XX
SQ Sequence 1266 BP; 216 A; 418 C; 390 G; 242 T; 0 other;

Alignment Scores:
Pred. No.: 4.9e-55 Length: 1266
Score: 1054.00 Matches: 184
Percent Similarity: 80.40% Conservative: 17
Best Local Similarity: 73.60% Mismatches: 49
Query Match: 73.89% Indels: 0
DB: 22 Gaps: 0

US-10-010-408-2 (1-250) x AAC97451 (1-1266)
Qy 1 MetArgGlySerProLeuIleHisLeuLeuAlaThrSerPheLeuCysLeuLeuSerMet 20
Db 10 ATGAGAGGCACACCGAAGACCCACCTCTCTGGCCCTTCTCCCTCTCTGCTCTCAAG 69
Qy 21 ValCysAlaGlnLeuCysArgThrProCysThrCysProThrThrProProGlnCysPro 40
Db 70 GTGCGTACCACAGTGTGCCCGCACACCATGTACCTGCCCTTGGCCACCTCCCGATGCCG 129
Qy 41 GlnGlyValProLeuValLeuAspGlyCysGlyCysLysValCysAlaArgArgLeu 60
Db 130 CTGGGAGTACCCCTGGTGCTGCTGATGGCTGTGGCTGCTGCCGGGTATGTGCACGGGGCTG 189
Qy 61 GlyGluSerCysAspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnPro 80
Db 190 GGGGAGCCCTGGGACCAACTCCACGTCTGCACGCCACGAGGCGCTGGTCTGCCAGCCC 249
Qy 81 GlyAlaGlyProGlyGlyHisGlyAlaValCysLeuLeuAspGluAspAspGlySerCys 100
Db 81 GlyAlaGlyProGlyGlyHisGlyAlaValCysLeuLeuAspGluAspAspGlySerCys 100
```



Qy 201 TrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGlnAsn 220  
Db 609 TGGGGACCCCTGCTGCACACCTGTGGGTGGGCATGGCCACCCGGGTGTCCACAGAAC 668  
Qy 221 ArgPheCysGlnLeuGluIleGlnArgLeuCysLeuProArgProCysLeuAlaAla 240  
Db 669 CGCTTCTGCCGACTGGAGACCCAGCGCGCTGTGCCTGTCCAGGCGCTGCCACCCCTCC 728  
Qy 241 ArgSerHisSerSerTrpAsnSerAlaPhe 250  
Db 729 AGGGTCCGACTCCACAAAACAGTGCCTTC 758

## RESULT 13

AAH76486

ID AAH76486 standard; DNA; 1293 BP.

XX

AC AAH76486;

XX

DT 06-AUG-1999 (first entry)

XX

DE Human WISP-2 protein nucleotide sequence SEQ ID NO:13.

XX

KW WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;  
connective tissue growth factor; cancer; melanoma; arteriosclerosis;  
leukaemia; lymphoid malignancy; haematopoiesis-related disorder;  
tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;  
kidney disorder; bone-related disorder; osteoporosis; trauma; burn;  
connective tissue disorder; catabolic state; inflammation;  
testicular-related disorder; angiogenesis; immunological disorder; ss.

XX

OS Homo sapiens.

XX

XX WO9921998-A1.

XX

PD 06-MAY-1999.

XX

XX 29-OCT-1998; 98WO-US22991.

XX

PR 14-APR-1998; 98US-0081695.

XX

PR 29-OCT-1997; 97US-0063704.

XX

PR 03-FEB-1998; 98US-0073612.

XX

XX (GETH ) GENENTECH INC.

XX

XX Botstein DA, Cohen RL, Goddard A, Gurney AL, Hillan K;

XX

XX Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WI;

XX

XX WPI; 1999-337420/28.

XX

XX P-PSDB; AAH17649.

XX

XX New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3

XX

XX Example 4; Page 174-175; 284pp; English.

XX

CC The present invention describes Wnt-1 induced secreted polypeptides,  
WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2  
and WISP-3 have homology to connective tissue growth factor (CTGF).  
CC products from the present invention can be used to treat WISP-related  
disorders such as breast, ovarian, and colon cancer or melanoma. The  
CC products can be used to treat arteriosclerosis. The products can also be  
used to treat other diseases e.g. benign and malignant tumours,  
CC leukaemia and lymphoid malignancies, neuronal, glial, astrocytic,  
CC hypothalamic and other glandular, macrophagal, epithelial, stromal, and  
CC blastocoele disorders, haematopoiesis-related disorders, tissue-growth  
disorders, skin disorders, desmoplasia, fibrotic lesions, kidney  
disorders, bone-related disorders such as osteoporosis, trauma such as  
burns, incisions, and other wounds, connective tissue disorders,  
catabolic states, testicular-related disorders, and inflammatory,  
CC angiogenic and immunologic disorders including arteriosclerosis. The  
products can also be used for detection and diagnosis especially of  
individuals with neoplastic cell growth or proliferation. The products  
can be used in the production of transgenic or knock-out animals.





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - nucleic search, using frame\_p2n model

Run on: August 1, 2003, 02:52:19 ; Search time 53 Seconds  
(without alignments)  
1446.589 Million cell updates/sec

Title: US-10-010-408-2  
Perfect score: 1440  
Sequence: 1 MRGSPHLHATSFLLCLSM.....LCIPPCLAARSHSNNSAP 250

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 441362 seqs, 15338381 residues  
Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Command line parameters:

-MODEL=framet.p2n.model -DEV=x1h  
-O=/cgn2\_1/USPTO.spool/US10010408/runat\_25072003\_101856\_17954/app\_query.fasta.1.391  
-DB=Issued\_Patents\_NA -OPMT=fastap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPEXT=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=bloms62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR\_SCORE=pcpt -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTEFT=ptlo -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USPR=US10010408 @CGN.1.1.40 @runat\_25072003\_101856\_17954 -NCPD=6 -ICPD=3  
-NO\_MMAP -IAREQUIDRY -NEG\_SCORES=0 -WAIT -DSFBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

## Database :

- Issued Patents\_NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCRTUS.COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfile1.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID | Description       |
|------------|--------|-------------|--------|-------|-------------------|
| 1          | 1308.5 | 90.9        | 1734   | 4     | US-09-182-145-17  |
| 2          | 1308.5 | 90.9        | 1734   | 4     | US-09-182-145-18  |
| 3          | 1064   | 73.9        | 1293   | 4     | US-09-182-145-13  |
| 4          | 1064   | 73.9        | 1293   | 4     | US-09-182-145-14  |
| 5          | 1047   | 72.7        | 738    | 4     | US-09-182-145-38  |
| 6          | 1045.5 | 72.6        | 841    | 4     | US-09-182-145-39  |
| 7          | 564    | 39.2        | 2075   | 1     | US-08-167-628-1   |
| 8          | 564    | 39.2        | 2075   | 1     | US-08-167-628-1   |
| 9          | 564    | 39.2        | 2075   | 1     | US-08-386-680-1   |
| 10         | 564    | 39.2        | 2075   | 1     | US-08-459-717-1   |
| 11         | 564    | 39.2        | 2075   | 1     | US-08-712-302-1   |
| 12         | 564    | 39.2        | 2075   | 2     | US-08-880-031-1   |
|            |        |             |        |       | Sequence 1, Appli |

|    |       |      |      |   |                   |                    |
|----|-------|------|------|---|-------------------|--------------------|
| 13 | 564   | 39.2 | 2075 | 4 | US-09-080-715-1   | Sequence 1, Appli  |
| 14 | 564   | 39.2 | 2075 | 4 | US-09-142-569-7   | Sequence 7, Appli  |
| 15 | 564   | 39.2 | 2075 | 5 | PCT-US96-08140-1  | Sequence 1, Appli  |
| 16 | 564   | 39.2 | 2998 | 3 | US-09-054-368-1   | Sequence 1, Appli  |
| 17 | 564   | 39.2 | 2998 | 3 | US-09-054-274-1   | Sequence 1, Appli  |
| 18 | 564   | 39.2 | 2998 | 4 | US-09-056-704-1   | Sequence 1, Appli  |
| 19 | 550.5 | 38.2 | 2267 | 4 | US-09-142-569-5   | Sequence 5, Appli  |
| 20 | 542.5 | 37.7 | 2350 | 4 | US-09-187-478-1   | Sequence 1, Appli  |
| 21 | 542.5 | 37.7 | 2350 | 4 | US-09-292-036-1   | Sequence 1, Appli  |
| 22 | 511   | 35.5 | 4234 | 4 | US-09-122-135-1   | Sequence 1, Appli  |
| 23 | 503.5 | 35.0 | 1766 | 4 | US-09-182-145-9   | Sequence 9, Appli  |
| 24 | 503.5 | 35.0 | 1766 | 4 | US-09-182-145-10  | Sequence 10, Appli |
| 25 | 501.5 | 34.8 | 1403 | 4 | US-09-182-145-23  | Sequence 23, Appli |
| 26 | 500.5 | 34.8 | 1480 | 4 | US-09-142-569-1   | Sequence 1, Appli  |
| 27 | 498.5 | 34.6 | 2830 | 4 | US-09-182-145-1   | Sequence 1, Appli  |
| 28 | 498.5 | 34.6 | 2830 | 4 | US-09-182-145-2   | Sequence 2, Appli  |
| 29 | 496.5 | 34.5 | 1418 | 4 | US-09-142-569-3   | Sequence 3, Appli  |
| 30 | 465   | 32.3 | 1128 | 2 | US-08-459-101A-1  | Sequence 1, Appli  |
| 31 | 446   | 31.0 | 1101 | 4 | US-09-182-145-29  | Sequence 29, Appli |
| 32 | 377.5 | 26.2 | 1335 | 4 | US-09-182-145-30  | Sequence 30, Appli |
| 33 | 377.5 | 26.2 | 1335 | 4 | US-09-182-145-31  | Sequence 31, Appli |
| 34 | 372.5 | 25.9 | 1232 | 4 | US-09-182-145-34  | Sequence 34, Appli |
| 35 | 372.5 | 25.9 | 1232 | 4 | US-09-182-145-35  | Sequence 35, Appli |
| 36 | 371.5 | 25.8 | 1142 | 4 | US-09-253-316-1   | Sequence 1, Appli  |
| 37 | 326.5 | 22.7 | 1062 | 4 | US-09-182-145-1   | Sequence 24, Appli |
| 38 | 220   | 15.3 | 693  | 4 | US-09-182-145-24  | Sequence 24, Appli |
| 39 | 220   | 15.3 | 1202 | 4 | US-09-182-145-26  | Sequence 26, Appli |
| 40 | 201.5 | 14.0 | 2541 | 2 | US-08-656-393-1   | Sequence 1, Appli  |
| 41 | 150   | 10.4 | 683  | 4 | US-09-182-145-25  | Sequence 25, Appli |
| 42 | 150   | 10.4 | 1183 | 4 | US-09-182-145-27  | Sequence 27, Appli |
| 43 | 139.5 | 9.7  | 8257 | 4 | US-09-484-9708-65 | Sequence 65, Appli |
| 44 | 139   | 9.7  | 1271 | 1 | US-08-464-339A-1  | Sequence 1, Appli  |
| 45 | 139   | 9.7  | 1271 | 5 | PCT-US94-14388-1  | Sequence 1, Appli  |

## ALIGNMENTS

RESULT 1  
US-09-182-145-17  
Sequence 17, Application US/09182145B  
Patent No. 6387657  
GENERAL INFORMATION:  
APPLICANT: Botstein, David A.  
APPLICANT: Cohen, Robert  
APPLICANT: Goddard, Audrey  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: Lawrence, David A.  
APPLICANT: Levine, Arnold J.  
APPLICANT: Pennica, Diane  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: P1176R2  
FILE REFERENCE: P1176R2  
CURRENT APPLICATION NUMBER: US/09/182,145B  
CURRENT FILING DATE: 1998-10-29  
EARLIER APPLICATION NUMBER: US 60/063,704  
EARLIER FILING DATE: 1997-10-29  
EARLIER APPLICATION NUMBER: US 60/073,612  
EARLIER FILING DATE: 1998-02-04  
EARLIER APPLICATION NUMBER: US 60/081,695  
EARLIER FILING DATE: 1998-04-14  
NUMBER OF SEQ ID NOS: 156  
SEQ ID NO 17  
LENGTH: 1734  
TYPE: DNA  
ORGANISM: Mus musculus  
US-09-182-145-17  
Alignment Scores:  
Pred. No.: 6e-95  
Score: 1308.50  
Percent Similarity: 93.63%  
Length: 1734  
Matches: 226  
Conservative: 9





Db 818 AACGATTCGCCAATCGAGATCCAGCGCTGTGTCTGTCCAGACCCCTGCCTGCA 759  
Qy 240 AAlarqserHisSerTrpAsnSerAlaPhe 250  
Db 758 TCCAGAGACCAAGCGCTCATGTGACAGATGCTTC 726

RESULT 3  
US-09-182-145-13  
; Sequence 13, Application US/09182145B  
; Patent No. 6387657  
; GENERAL INFORMATION:  
; APPLICANT: Botstein, David A.  
; APPLICANT: Cohen, Robert  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth J.  
; APPLICANT: Lawrence, David A.  
; APPLICANT: Levine, Arnold J.  
; APPLICANT: Pennica, Diane  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Wood, William I.  
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: P1176R2  
; CURRENT APPLICATION NUMBER: US/09/182,145B  
; EARLIER FILING DATE: 1998-10-29  
; EARLIER APPLICATION NUMBER: US 60/063,704  
; EARLIER FILING DATE: 1997-10-29  
; EARLIER APPLICATION NUMBER: US 60/073,612  
; EARLIER FILING DATE: 1998-02-04  
; EARLIER APPLICATION NUMBER: US 60/081,695  
; EARLIER FILING DATE: 1998-04-14  
; NUMBER OF SEQ ID NOS: 156  
; SEQ ID NO 13  
; LENGTH: 1293  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-182-145-13

Alignment Scores:  
Pred. No.: 9,63e-76 Length: 1293  
Score: 1064.00 Matches: 184  
Percent Similarity: 80.40% Conserved: 17  
Best Local Similarity: 73.60% Mismatches: 49  
Query Match: 73.89% Indels: 0  
Gaps: 0

US-10-010-408-2 (1-250) x US-09-182-145-13 (1-1293)

Qy 1 MetaArgLysSerProLeuIleHisLeuAlaThrSerPheLeuCysLeuSerMet 20  
Db 22 ATGAGAGGACACAGGAGACCCAGCTCTGCTGCTCTCTCTCTCTCTCTCTCAAG 81  
Qy 21 ValCysAlaGlnLeuCysArgThrProCysThrCysProThrProGlnCysPro 40  
Db 82 GTGGGTACCCAGCTGTGCGGACACATGACCTGCGGACCTCCCGCATGCGCG 141  
Qy 41 GlnGlyValProLeuValLeuAspGlyCysGlyCysGlyCysValCysAlaArgArgLeu 60  
Db 142 CTGGGATGTACCCCTGGGTGTGATGCTGTGCTGTGCTGGGATGTGTGACGCGGCTG 201  
Qy 61 GlyLysSerCysAspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnPro 80  
Db 202 GGGGAGCGCCGTGCGACCACTCCAGCTGTGCGACCCAGCCAGCGCTGTGTGCGACCCC 261  
Qy 81 GlyAlaGlyProGlyGlyHisGlyAlaValCysLeuLeuAspGlnAspAspGlySerCys 100  
Db 262 GGGGACAGACCCGGTGTGCGGCGGCGCTGTGCTGTGCGACAGAGGACGACAGACCTGT 321  
Qy 101 GlnValAsnGlyArgArgGlyTrpLeuAspGlyGlnThrPheLeuProAsnGlyValLeu 120  
Db 322 GAGGTGAAGCGCCGCTGTATCGGGAAGGAGACCTTCCAGCCCACTGCGACGATCGCG 381  
Qy 121 CysArgCysAspAspGlyGlyPheThrCysLeuProLeuCysSerGlnAspValArgLeu 140

Db 382 TGGCGGTGCGAGAGCGCGGCTTCACCTGCTGCGCGCTGTGACAGCGAGATGTGGGCTG 441  
Qy 141 ProSerTrpAspCysProAlaGlyProGlyArgIleGlnValProGlyLysCysCysProGlu 160  
Db 442 CCCAGCTGGAGATGCCCCCAGCCAGGAGGAGGTGCTGTGGCAAGTGTGCTGCGGAG 501  
Qy 161 TrpValCysAspGlnGlyValThrProAlaIleGlnArgSerThrAlaGlnGlyHisGln 180  
Db 502 TGGGTGTGCGGCGCCAGAGAGGGGACTGGGACCCAGCCCTTCAGGCCCAAGACCCAG 561  
Qy 181 LeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThrAla 200  
Db 562 TTTTCTGCGCTTGTGCTGTCTGCTGCGCGCCCTGTGTGCTGCTGCGCGAGATGTGAGCGGCC 621  
Qy 201 TrpGlyProCysSerThrTrpCysGlyLeuGlyIleAlaThrArgValSerAsnGlnAsn 220  
Db 622 TGGGACCTGTGCTGACACCTGTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 681  
Qy 221 ArgPheCysGlnLeuGlnIleGlnArgArgLeuCysLeuProArgProCysLeuAlaAla 240  
Db 682 CGCTTGTGCGGACTGTGAGACCCAGCGCGCTGTGCTGTGCTGTGCGAGGCTTGCACCTCC 741

Qy 241 ArgSerHisSerSerTrpAsnSerAlaPhe 250  
Db 742 AGGGGTGCGAGTCCACAAACAGTGCCTTC 771

RESULT 4  
US-09-182-145-14/c  
; Sequence 14, Application US/09182145B  
; Patent No. 6387657  
; GENERAL INFORMATION:  
; APPLICANT: Botstein, David A.  
; APPLICANT: Cohen, Robert  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth J.  
; APPLICANT: Lawrence, David A.  
; APPLICANT: Levine, Arnold J.  
; APPLICANT: Pennica, Diane  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Wood, William I.  
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: P1176R2  
; CURRENT APPLICATION NUMBER: US/09/182,145B  
; EARLIER FILING DATE: 1998-10-29  
; EARLIER APPLICATION NUMBER: US 60/063,704  
; EARLIER FILING DATE: 1997-10-29  
; EARLIER APPLICATION NUMBER: US 60/073,612  
; EARLIER FILING DATE: 1998-02-04  
; EARLIER APPLICATION NUMBER: US 60/081,695  
; EARLIER FILING DATE: 1998-04-14  
; NUMBER OF SEQ ID NOS: 156  
; SEQ ID NO 14  
; LENGTH: 1293  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-182-145-14

Alignment Scores:  
Pred. No.: 9,63e-76 Length: 1293  
Score: 1064.00 Matches: 184  
Percent Similarity: 80.40% Conserved: 17  
Best Local Similarity: 73.60% Mismatches: 49  
Query Match: 73.89% Indels: 0  
Gaps: 0

US-10-010-408-2 (1-250) x US-09-182-145-14 (1-1293)

Qy 1 MetaArgLysSerProLeuIleHisLeuAlaThrSerPheLeuCysLeuSerMet 20  
Db 1272 ATGAGAGGACACAGGAGACCCAGCTCTGCTGCTCTCTCTCTCTCTCTCAAG 1213







```

Db      580 CCGAGAGAGGTCAAGCTCCCGGAAATGCTGCGAGAGTGGTGTGAGAG----- 633
Qy      168 ThrProAlaIleGlnArgSerThrAlaGlnIlyHisGlnLeuSerAlaLeu----- 184
Db      634 ---CCCAAGACCAA-----ACCGTGGTGGGCTGCGCGGGCTTACCGACGTGAA 684
Qy      185 -----ValThrProAlaSerAlaSpAlaProCysProAsnTrpSerThrAla 200
Db      685 GACACGTTGGCCCAAGACCACTATGATTAGAGCAACTGCTGGTCCAGACCAAG 744
Qy      201 TrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGlnAsn 220
Db      745 TGGAGCGGCTGTCCAGACCTGTGGGATGGGCAATCTCCACCGGGTTACCAATGACAAAC 804
Qy      221 ArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProAlaArgProCysLeuAla 239
Db      805 GCCTCTGACAGGCTAGAGAACAGAGCCGCTGTGCATGTGTCAGAGCCCTTGGGAAGCT 861

RESULT 9
US-08-459-717-1
: Sequence 1, Application US/08459717
: Patent No. 5770209
: GENERAL INFORMATION:
: APPLICANT: Grotendorst, Gary R.
: APPLICANT: Bradham Jr., Douglas M.,
: TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR
: NUMBER OF SEQUENCES: 2
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Spensley Horn Jubas & Lubitz
: STREET: 4225 Executive Square, Suite 1400
: CITY: La Jolla
: STATE: CA
: COUNTRY: US
: ZIP: 92037
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/459,717
: FILING DATE: 02-JUN-1995
: CLASSIFICATION: 536
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/752,427
: FILING DATE: 30-AUG-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Wetherell, Jr. Ph.D., John W.
: REGISTRATION NUMBER: 31,678
: REFERENCE/DOCKET NUMBER: PD-1294
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 619-455-5100
: TELEFAX: 619-455-5110
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2075 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: IMMEDIATE SOURCE:
: CLONE: DB60R32
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 130..1177
: US-08-459-717-1

Alignment Scores:
Pred. No.: 7,41e-36 Length: 2075
Score: 564.00 Matches: 112
Percent Similarity: 58.58% Conservative: 28
Best Local Similarity: 46.86% Mismatches: 81

```

```

Query Match: 39.17% Indels: 18
DB: 1 Gaps: 6

US-10-010-408-2 (1-250) x US-08-459-717-1 (1-2075)
Qy      13 SerPheLeuCysLeuLeuSerMet-----ValCysAlaGlnLeuCysArgThr 28
Db      163 GCCTTCGTGTCCTCTCTGCGCCCTCGACGCCGGCGCGCGGCGGCAAGAACTGCAGCGCG 222
Qy      29 ProCysThrCysProTrpThrPro---ProGlnCysProGlnGlyValProLeuValLeu 47
Db      223 CCGTCCCGGCTCCCGGAGACCGCGCGCGCGCTCCCGCGGGGCGTGAAGCTGTGCTGTG 282
Qy      48 AspGlyCysGlyCysCysValCysAlaArgArgLeuGlyIlySerCysAspHisLeu 67
Db      283 GACGGCTGGCGCTGCTGCGCGCTGCGCCCAAGACAGCTGGCGGCGAGCTGTGCACCGAGCGC 342
Qy      68 HisValCysAspProSerGlnGlyLeuValCysGlnProGlnAlaGlyProGlyGlyHis 87
Db      343 GACCCCTGCGACCCGCAACAGGCGCTTCTGTGACTTGCAGCTCCCGCGCCACCGCAAG 402
Qy      88 GlyAlaValCysLeuLeuAspGlnAspAspGlySerCysGlyValAsnGlyArgArgTyr 107
Db      403 ATCGGCGGTGCACCGCC---AAAGATGTGCTCCCTCCATCTTGGGTGTACGGGTGTAC 459
Qy      108 LeuAspGlyIlyThrPheLeuProAsnCysArgValLeuCysArgCysAspAspGlyGly 127
Db      460 CGCAGCGAGAGTCTTCACAGACAGCTGCAGAGTACAGTGCACGTGCTGAGCGGGCGG 519
Qy      128 PheThrCysLeuProLeuCysSerGlnAspValArgLeuProSerTrpAspCysProArg 147
Db      520 GTGGGCTGTGATGCCCTGTGACGATGACGTGTGTCGCCAGCGCTGACGCCCTTC 579
Qy      148 ProLysArgIleGlnValProGlyIlyCysCysProGluTrpValCysAspGlnGlyVal 167
Db      580 CCGAGGAGGTCAAGCTGCCCGGAAAGTGCAGAGGTGGGTGTGAGAGAG----- 633
Qy      168 ThrProAlaIleGlnArgSerThrAlaGlnIlyHisGlnLeuSerAlaLeu----- 184
Db      634 ---CCCAAGACCAA-----ACCGTGGTGGGCTGCGCTCCGCGCTTACCGACGTGAA 684
Qy      185 -----ValThrProAlaSerAlaSpAlaProCysProAsnTrpSerThrAla 200
Db      685 GACACGTTGGCCCAAGACCACTATGATTAGAGCAACTGCTGGTCCAGACCAAG 744
Qy      201 TrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGlnAsn 220
Db      745 TGGAGCGGCTGTCCAGACCTGTGGGATGGGCAATCTCCACCGGGTTACCAATGACAAAC 804
Qy      221 ArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProAlaArgProCysLeuAla 239
Db      805 GCCTCTGACAGGCTAGAGAACAGAGCCGCTGTGCATGTGTCAGAGCCCTTGGGAAGCT 861

RESULT 10
US-08-712-302-1
: Sequence 1, Application US/08712302
: Patent No. 5763187
: GENERAL INFORMATION:
: APPLICANT: Grotendorst, Gary R.
: APPLICANT: Bradham Jr., Douglas M.,
: TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR
: NUMBER OF SEQUENCES: 2
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Spensley Horn Jubas & Lubitz
: STREET: 4225 Executive Square, Suite 1400
: CITY: La Jolla
: STATE: CA
: COUNTRY: US
: ZIP: 92037
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/712,302
FILING DATE: 11-SEP-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/386,680
FILING DATE: 10-FEB-1995
APPLICATION NUMBER: US/08/167,628
FILING DATE:
APPLICATION NUMBER: US/07/752,427
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Wetherell, Jr. Ph.D., John W.
REGISTRATION NUMBER: 31,678
REFERENCE/DOCKET NUMBER: PD-1294
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-455-5100.
TELEFAX: 619-455-5110
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2075 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
IMMEDIATE SOURCE:
CLONE: DB60R32
FEATURE:
NAME/KEY: CDS
LOCATION: 130..1177
US-08-712-302-1

```

## Alignment Scores:

```

Pred. No.: 7.41e-36      Length: 2075
Score: 564.00           Matches: 112
Percent Similarity: 58.588      Conservative: 28
Best Local Similarity: 46.868    Mismatches: 81
Query Match: 39.17%           Indels: 18
DB: 1                      Gaps: 6

```

US-10-010-408-2 (1-250) x US-08-712-302-1 (1-2075)

```

QY 13 SerPheLeuCYsLeuLeuSerMet-----ValCYsAlaGlnLeuCYsArgThr 28
DB 163 GCCTTCGTGGTCTCTCGCCCTTCGACGCGCGCGCGGTGGCCAGCAACTGACGCGG 222
QY 29 ProCYsThrCYsProTirPhePro---ProGlnCYsProGlnGlyValProLeuValLeu 47
DB 223 CCGTCCCGGTGCCGAGACGCGCGCGCGCGCTCCGCGCGGTGAGCCTCGGTG 282
QY 48 AspGlyCYsGlyCYsCYsLYsValCYsAlaArgLeuGlyGlySerCYsAspHisLeu 67
DB 283 GACGGCTGGGTCTGCTGCGCGCTTCGCGCAAGACAGCTGGGAGACTGTCACCGACGCC 342
QY 68 HisValCYsAspProSerGlnGlyLeuValCYsGlnProGlyAlaGlyProGlyGlyHis 87
DB 343 GACCCCTGCGACCCACACAAAGGCGCTTCTGTGACTTCCCGCGCGCAACGCAAG 402
QY 88 GlyAlaValCYsLeuLeuAspGlnAspAspGlySerCYsGlnValAlaGlnArgArgTyr 107
DB 403 ATGGCGGTGTGACCGCC---AAAGATGGTGTCTCCCTGATCTTGGGTGTACGGTGTAC 459
QY 108 LeuAspGlyGlnThrPhelysProAsnCYsArgValLeuCYsArgCYsAspAspGlyGly 127
DB 460 CGCAGCGGAGATCTCTTCAGAGACAGCTCAAGTACAGTACAGTGGCTGACGCGGGG 519
QY 128 PheThrCYsLeuProLeuCYsSerGlnAspValArgLeuProSerTirPaspCYsProArg 147
DB 520 GTGGGCTGATCCCTGTCAGACATGACGTTGCTGCTCCACGCGCTGACGCGCTTC 579
QY 148 ProLYsArgGlnLeuValProGlyLYsCYsProGlnThrValCYsAspGlnGlyVal 167

```

```

DB 580 CCGAGAGAGGTCACAGTCCCGCGGAATGCTCCGAGAGATGGGTGTGACAG----- 633
QY 168 ThrProAlaIleGlnArgSerThrAlaGlnGlyHisGlnLeuSerAlaLeu----- 184
DB 634 ---CCCAAGGACCA-----ACCGGTGGGCGCTGCCCTCGCGGCTTACCGACTGGA 684
QY 185 -----ValThrProAlaSerAlaAspAlaProCYsProAsnTirPserThrAla 200
DB 685 GACAGGTTGGCCGACAGCCCACTATGATTAGAGCCACTGCTGTCAGACACAGAG 744
QY 201 TTPGlyProCYsSerThrCYsGlyLeuGlyIleAlaThrArgValSerAsnGlnAsn 220
DB 745 TGGAGCGCTGTTCACAGACCTGTGGAGGCGCATCTCCACCGGCTTACCAATGACAC 804
QY *221 ArgPheCYsGlnLeuGlnIleGlnArgLeuCYsLeuProArgProCYsLeuAla 239
DB 805 GCCTCCTGACGCTAGAGAGACAGCGCCGCTGTGATGTGATGATGATGATGATGATGAT 861

```

## RESULT 11

US-08-880-031-1

Sequence 1, Application US/08880031

Patent No. 5916756

GENERAL INFORMATION:

APPLICANT: Grotendorst, Gary R.

APPLICANT: Bradham Jr., Douglas M.,

TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: Spensley Horn Judas &amp; Lubitz

STREET: 4225 Executive Square, Suite 1400

CITY: La Jolla

STATE: CA

COUNTRY: US

ZIP: 92037

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/880,031

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/167,628

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Wetherell, Jr. Ph.D., John W.

REGISTRATION NUMBER: 31,678

REFERENCE/DOCKET NUMBER: PD-1294

TELECOMMUNICATION INFORMATION:

TELEPHONE: 619-455-5100

TELEFAX: 619-455-5110

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 2075 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

IMMEDIATE SOURCE:

CLONE: DB60R32

FEATURE:

NAME/KEY: CDS

LOCATION: 130..1177

US-08-880-031-1

## Alignment Scores:

```

Pred. No.: 7.41e-36      Length: 2075
Score: 564.00           Matches: 112
Percent Similarity: 58.588      Conservative: 28
Best Local Similarity: 46.868    Mismatches: 81
Query Match: 39.17%           Indels: 18

```



QY 168 ThrProAlaIleGlnArgSerThrAlaGlnGlyHisGlnLeuSerAlaLeu----- 184  
 DB 634 ---CCCAAGACCAA-----ACCGTGGTGGCGCTGGCGGTACCGAGCTGGA 684  
 QY 185 -----ValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThrAla 200  
 DB 685 GACACGTTGGCCAGACCACTATGATTAGAGCCACTGCTGCTGCAGACCAAGAG 744  
 QY 201 TrpGlyProCysSerThrThrcysGlyLeuGlyIleAlaThrArgValSerAsnGlnAsn 220  
 DB 745 TGGACCGCCTGTTCGAAGACCTGTGGATGGCATCTCCACCCGGGTTACCAATGACAA 804  
 QY 221 ArgPheCysGlnLeuGlnIleGlnArgArgLeuCysLeuProArgProCysLeuAla 239  
 DB 805 GCCTCCTGACGCTAGAGAGACGAGCGCGCTGCTGATGTCAGCGCTTGCAGACT 861  
 RESULT 13  
 US-09-080-715-1  
 ; Sequence 1, Application US/09080715  
 ; Patent No. 6190884  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Grotendorst, Gary R.  
 ; APPLICANT: Bradham Jr., Douglas M.,  
 ; TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR  
 ; NUMBER OF SEQUENCES: 2  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Spensley Horn Judas & Lubitz  
 ; STREET: 4225 Executive Square, Suite 1400  
 ; CITY: La Jolla  
 ; STATE: CA  
 ; COUNTRY: US  
 ; ZIP: 92037  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/080,715  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/167,628  
 ; FILING DATE:  
 ; APPLICATION NUMBER: US/07/752,427  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Wetherell, Jr. Ph.D., John W.  
 ; REGISTRATION NUMBER: 31,678  
 ; REFERENCE/DOCKET NUMBER: PD-1294  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 619-455-5100  
 ; TELEFAX: 619-455-5110  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 2075 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: cDNA  
 ; IMMEDIATE SOURCE:  
 ; CLONE: DB60R32  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 130..1177  
 ; US-09-080-715-1  
 Alignment Scores:  
 Pred. No.: 7.41e-36 Length: 2075  
 Score: 564.00 Matches: 112  
 Percent Similarity: 58.58 Conservative: 28  
 Best Local Similarity: 46.86 Mismatches: 81

Query Match: 39.17% Indels: 18  
 DB: 4 Gaps: 6  
 US-10-010-408-2 (1-250) x US-09-080-715-1 (1-2075)  
 QY .13 SerPheLeuCysLeuLeuSerMet-----ValCysAlaGlnLeuCysArgThr 28  
 DB 163 GCGTTCGCTGCTCCCTCCCTCCCTGACACCGCGCGCGCGCGCGCGCGCGCG 222  
 QY 29 ProCysThrCysProTrpThrPro---ProGlnCysProGlnGlyValProLeuValLeu 47  
 DB 223 CCGTCCCGGTGCGCCGAGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 282  
 QY 48 AspGlyCysGlyCysCysLysValCysAlaArgArgLeuGlyLysCysAspHisLeu 67  
 DB 283 GAGCGCTGGCGTGTGCTGCGCGCTGCGCGCAAGCTGCGGAGACTGTGCACAGCGC 342  
 QY 68 HisValCysAspProSerGlnGlyLeuValCysGlnProGlyAlaGlyProGlyLys 87  
 DB 343 GACCCCTGGACCCGACACAGGCGCTCTTCTGTGACTTGTGCTCCCGCCGCAAG 402  
 QY 88 GlyAlaValCysLeuLeuAspGlnAspAspGlySerCysGluValAsnGlyArgArgTyr 107  
 DB 403 ATCGCGCTGTGACCGCC---AAAGATGCTGCTCCCTGATCTTGGTGTGCGGTGAC 459  
 QY 108 LeuAspGlyGluThrPheLysProAsnCysArgValLeuCysArgCysAspAspGly 127  
 DB 460 CCGACGCGAGAGTCTTCACAGACAGCTGCAAGTACAGTGCAGCTGCGCGCGCG 519  
 QY 128 PheThrCysLeuProLeuCysSerGlnAspValArgLeuProSerTrpAspCysProArg 147  
 DB 520 GTGGGCTGATGCCCCCTGTGACATGACGCTTGTGCTGCTGCGCGCGCGCGCTTC 579  
 QY 148 ProLysArgIleGlnValProGlyLysCysCysProGluTrpValCysAspGlnGlyVal 167  
 DB 580 CCGAGAGAGGTAAAGCTGCGCGGAAATCTCGAGAGTGGTGTGTGACAG----- 633  
 QY 168 ThrProAlaIleGlnArgSerThrAlaGlnGlyHisGlnLeuSerAlaLeu----- 184  
 DB 634 ---CCCAAGACCAA-----ACCGTGGTGGCGCTGGCGGTACCGAGCTGGA 684  
 QY 185 -----ValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThrAla 200  
 DB 685 GACACGTTGGCCAGACCACTATGATTAGAGCCACTGCTGCTGCAGACCAAGAG 744  
 QY 201 TrpGlyProCysSerThrThrcysGlyLeuGlyIleAlaThrArgValSerAsnGlnAsn 220  
 DB 745 TGGACCGCCTGTTCGAAGACCTGTGGATGGCATCTCCACCCGGGTTACCAATGACAA 804  
 QY 221 ArgPheCysGlnLeuGlnIleGlnArgArgLeuCysLeuProArgProCysLeuAla 239  
 DB 805 GCCTCCTGACGCTAGAGAGACGAGCGCGCTGCTGATGTCAGCGCTTGCAGACT 861  
 RESULT 14  
 US-09-142-569-7  
 ; Sequence 7, Application US/09142569  
 ; Patent No. 6413735  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lau, Lester F.  
 ; TITLE OF INVENTION: Extracellular Matrix Signalling Molecules  
 ; NUMBER OF SEQUENCES: 17  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
 ; STREET: 6300 Sears Tower, 233 South Wacker Drive  
 ; CITY: Chicago  
 ; STATE: Illinois  
 ; COUNTRY: United States of America  
 ; ZIP: 60606-6402  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.30



```

CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/142,566
: FILING DATE: 02-Apr-1999
: CLASSIFICATION: <Unknown>
: ATTORNEY/AGENT INFORMATION:
:   NAME: Clough, David W.
:   REGISTRATION NUMBER: 36,107
:   REFERENCE/DOCKET NUMBER: 28758/33766
: TELECOMMUNICATION INFORMATION:
:   TELEPHONE: 312/474-6300
:   TELEFAX: 312/474-0448
:   TELEX: 25-3856
: INFORMATION FOR SEQ ID NO: 7:
:   SEQUENCE CHARACTERISTICS:
:     LENGTH: 2075 base pairs
:     TYPE: nucleic acid
:     STRANDEDNESS: single
:     TOPOLOGY: linear
: MOLECULE TYPE: DNA
: FEATURE:
:   NAME/KEY: misc feature
:   OTHER INFORMATION: "CTGF cDNA coding sequence"
: SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-142-569-7

Alignment Scores:
Pred. No.: 7,41e-36      Length: 2075
Score: 564.00           Matches: 112
Percent Similarity: 58.58%      Conservative: 28
Best Local Similarity: 46.86%   Mismatches: 81
Query Match: 39.17%           Indels: 18
DB: 4                    Gaps: 6

US-10-010-408-2 (1-250) x US-09-142-569-7 (1-2075)
QY 13 SerPheLeuCySLeuSerMet-----ValCysAlaGlnLeuCysArgThr 28
Db 163 GCCTTCGTGCTCTCTCGCCCTCTGCAGCCGCCGCCCTCGGCCAGAACTGCAGCGCG 222
QY 29 ProCysThrCysProTPrthPro---ProGlnCysProGlnGlyValProLeuValLeu 47
Db 223 CCGTGCCTGCTCCCGGAGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 282
QY 48 AspGlyCysGlyCysCysLysValCysAlaIarGlarGleuGlyGluSerCysAspHisLeu 67
Db 283 GACGGCTGCGCTCTGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 342
QY 68 HisValCysAspProSerGlnGlyLeuValCysGlnProGlyAlaGlyProGlyGlyHis 87
Db 343 GACCCGTGCGACCGCGCAAGAGGCGCTCTGTGTGACTTGTCCCGCGCGCAACCGCGAAG 402
QY 88 GlyAlaValCysLeuLeuAspGluAspAspGlySerCysGluValAsnGlyArgArgTyr 107
Db 403 ATCGCGCTGTGCACCGCC--AAAGATGTGTCTCTCCCTGCATCTTGTGGGTATCGGTATC 459
QY 108 LeuAspGlyGluThrPheLysProAsnCysArgValLeuCysArgCysAspAspGlyCly 127
Db 460 CGCAGCGGAGAGTCTCTTCCAGACGAGCTGCAAGTACCAAGTACAGTGTGCTGAGCGGCG 519
QY 128 PheThrCysLeuProLeuCysSerGlnAspValArgLeuProSerTyrPaspCysProArg 147
Db 520 GTGGGCTGCATGCCCTGTGCAGCATGTGCTGTGCGCACCGCTGACTGCGCCCTTC 579
QY 148 ProLysArgGlnGlnValProGlyLysCysCysProGluTyrValCysAspGlnGlyVal 167
Db 580 CCGAGGAGGCTCAAGTCCCGGGAAGATCTGCAGAGGAGTGTGTGTGAGAGAG----- 633
QY 168 ThrProAlaIleGlnArgSerThrAlaGlnGlyHisGlnLeuSerAlaLeu----- 184
Db 634 ---CCCAAGACCA-----ACCGTGTGTGGGCTGCGCTGCGGCGCTTACCAGTGA 684
QY 185 -----ValThrProAlaSerAlaAspAlaProCysProAsnTyrSerThrAla 200

```

```

Db 685 GACACGTTTGGCCCAAGCCCAACTATGATTAGACCAACTGCTGTCCAGACCAAGAG 744
QY 201 TrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGlnAsn 220
Db 745 TGGAGCGGCTGTTCAGAGACTGTGGATGGGCAATTCACACCGGGTTACCATAAGACAAAC 804
QY 221 ArgPheCysGlnLeuGluIleGlnArgLeuArgLeuCysLeuProArgProCysLeuAla 239
Db 805 GCCTTCGTGACGAGCTATGAGAACGACAGCGCGCTGTGATGTGTACAGCGCTTGGGAAGCT 861

RESULT 15
PCT-0596-08140-1
: Sequence 1, Application PC/TUS9608140
: GENERAL INFORMATION:
: APPLICANT: University of South Florida
: TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR
: NUMBER OF SEQUENCES: 2
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: FISH & RICHARDSON P.C.
: STREET: 4225 Executive Square, Suite 1400
: CITY: La Jolla
: STATE: CA
: COUNTRY: US
: ZIP: 92037
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US96/08140
: FILING DATE: 30-MAY-1996
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
:   NAME: Haile, Ph.D., Lisa A.
:   REGISTRATION NUMBER: 38,347
:   REFERENCE/DOCKET NUMBER: 07414/003W01
: TELECOMMUNICATION INFORMATION:
:   TELEPHONE: 619-678-5070
:   TELEFAX: 619-678-5099
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
:   LENGTH: 2075 base pairs
:   TYPE: nucleic acid
:   STRANDEDNESS: single
:   TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: IMMEDIATE SOURCE:
: CLONE: CTGF
: FEATURE:
:   NAME/KEY: CDS
:   LOCATION: 130..1177
: PCT-0596-08140-1

US-10-010-408-2 (1-250) x PCT-0596-08140-1 (1-2075)
QY 13 SerPheLeuCySLeuSerMet-----ValCysAlaGlnLeuCysArgThr 28
Db 163 GCCTTCGTGCTCTCTCGCCCTCTGCAGCCGCCGCCCTCGGCCAGAACTGCAGCGCG 222
QY 29 ProCysThrCysProTPrthPro---ProGlnCysProGlnGlyValProLeuValLeu 47
Db 223 CCGTGCCTGCTCCCGGAGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 282
QY 48 AspGlyCysGlyCysCysLysValCysAlaIarGlarGleuGlyGluSerCysAspHisLeu 67

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: August 1, 2003, 03:33:55 ; Search time 231 Seconds

(Without alignments)  
2232.685 Million cell updates/sec

Title: US-10-010-408-2

Perfect score: 1440

Sequence: 1 MRGSPHLHLATSFLLLSM.....LCIPRCLARSHSSWNSAF 250

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1439767 segs, 1031500376 residues

Total number of hits satisfying chosen parameters: 2879534

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=framed.p2n.model -DEV=x1h  
-Q=/cgn2\_1/USFTO.spool/US10010408/runat\_25072003\_101856\_17969/app\_query.fasta.1.391  
-DB=Published.Applications.NA -OFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=bl0sum62  
-TRANS=human40.ccd1 -LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US10010408 @CCN.1.1.57 @runat\_25072003\_101856\_17969  
-NCPU=6 -ICPU=3 -NO\_MMAP -LARGESUBJECT -NEG\_SCORES=0 -WAIT -DSBLOCK=100  
-LONGLOS -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications.NA.\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq1.\*  
11: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq2.\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq3.\*  
13: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
15: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID              | Description       |
|------------|-------|-------------|--------|-----------------|-------------------|
| 1          | 1440  | 100.0       | 753    | US-10-010-408-3 | Sequence 3, Appl1 |

|  | 2  | 1440   | 100.0 | 1708 | 15  | US-10-010-408-1    | Sequence 1, Appl1 |
|--|----|--------|-------|------|---|--------------------|-------------------|
|  | 2  | 1323   | 91.9  | 681  | 15 <th>US-10-010-408-12</th> <th>Sequence 12, Appl</th>   | US-10-010-408-12   | Sequence 12, Appl |
|  | 4  | 1308.5 | 90.9  | 1734 | 15 <th>US-10-112-267-17</th> <th>Sequence 17, Appl</th>   | US-10-112-267-17   | Sequence 17, Appl |
|  | 5  | 1308.5 | 90.9  | 1734 | 15 <th>US-10-112-267-18</th> <th>Sequence 18, Appl</th>   | US-10-112-267-18   | Sequence 18, Appl |
|  | 6  | 1064   | 73.9  | 1266 | 14 <th>US-10-137-866-319</th> <th>Sequence 319, App</th>  | US-10-137-866-319  | Sequence 319, App |
|  | 7  | 1064   | 73.9  | 1266 | 14 <th>US-10-146-727-319</th> <th>Sequence 319, App</th>  | US-10-146-727-319  | Sequence 319, App |
|  | 8  | 1064   | 73.9  | 1266 | 14 <th>US-10-146-788-319</th> <th>Sequence 319, App</th>  | US-10-146-788-319  | Sequence 319, App |
|  | 9  | 1064   | 73.9  | 1266 | 14 <th>US-10-152-880-319</th> <th>Sequence 319, App</th>  | US-10-152-880-319  | Sequence 319, App |
|  | 10 | 1064   | 73.9  | 1266 | 14 <th>US-10-153-934-319</th> <th>Sequence 319, App</th>  | US-10-153-934-319  | Sequence 319, App |
|  | 11 | 1064   | 73.9  | 1266 | 14 <th>US-10-028-072-319</th> <th>Sequence 319, App</th>  | US-10-028-072-319  | Sequence 319, App |
|  | 12 | 1064   | 73.9  | 1266 | 15 <th>US-10-121-049-319</th> <th>Sequence 319, App</th>  | US-10-121-049-319  | Sequence 319, App |
|  | 13 | 1064   | 73.9  | 1266 | 15 <th>US-10-123-904-319</th> <th>Sequence 319, App</th>  | US-10-123-904-319  | Sequence 319, App |
|  | 14 | 1064   | 73.9  | 1266 | 15 <th>US-10-140-470-319</th> <th>Sequence 319, App</th>  | US-10-140-470-319  | Sequence 319, App |
|  | 15 | 1064   | 73.9  | 1266 | 15 <th>US-10-140-474-319</th> <th>Sequence 319, App</th>  | US-10-140-474-319  | Sequence 319, App |
|  | 16 | 1064   | 73.9  | 1266 | 15 <th>US-10-175-446-319</th> <th>Sequence 319, App</th>  | US-10-175-446-319  | Sequence 319, App |
|  | 17 | 1064   | 73.9  | 1266 | 15 <th>US-10-176-918-319</th> <th>Sequence 319, App</th>  | US-10-176-918-319  | Sequence 319, App |
|  | 18 | 1064   | 73.9  | 1266 | 15 <th>US-10-176-921-319</th> <th>Sequence 319, App</th>  | US-10-176-921-319  | Sequence 319, App |
|  | 19 | 1064   | 73.9  | 1266 | 15 <th>US-10-137-865-319</th> <th>Sequence 319, App</th>  | US-10-137-865-319  | Sequence 319, App |
|  | 20 | 1064   | 73.9  | 1266 | 15 <th>US-10-140-474-319</th> <th>Sequence 319, App</th>  | US-10-140-474-319  | Sequence 319, App |
|  | 21 | 1064   | 73.9  | 1266 | 15 <th>US-10-142-431-319</th> <th>Sequence 319, App</th>  | US-10-142-431-319  | Sequence 319, App |
|  | 22 | 1064   | 73.9  | 1266 | 15 <th>US-10-143-114-319</th> <th>Sequence 319, App</th>  | US-10-143-114-319  | Sequence 319, App |
|  | 23 | 1064   | 73.9  | 1266 | 15 <th>US-10-140-002-319</th> <th>Sequence 319, App</th>  | US-10-140-002-319  | Sequence 319, App |
|  | 24 | 1064   | 73.9  | 1266 | 15 <th>US-10-142-419-319</th> <th>Sequence 319, App</th>  | US-10-142-419-319  | Sequence 319, App |
|  | 25 | 1064   | 73.9  | 1266 | 15 <th>US-10-123-262-319</th> <th>Sequence 319, App</th>  | US-10-123-262-319  | Sequence 319, App |
|  | 26 | 1064   | 73.9  | 1266 | 15 <th>US-10-142-423-319</th> <th>Sequence 319, App</th>  | US-10-142-423-319  | Sequence 319, App |
|  | 27 | 1064   | 73.9  | 1266 | 15 <th>US-10-121-050-319</th> <th>Sequence 319, App</th>  | US-10-121-050-319  | Sequence 319, App |
|  | 28 | 1064   | 73.9  | 1266 | 15 <th>US-10-141-755-319</th> <th>Sequence 319, App</th>  | US-10-141-755-319  | Sequence 319, App |
|  | 29 | 1064   | 73.9  | 1266 | 15 <th>US-10-143-032-319</th> <th>Sequence 319, App</th>  | US-10-143-032-319  | Sequence 319, App |
|  | 30 | 1064   | 73.9  | 1266 | 15 <th>US-10-123-108-319</th> <th>Sequence 319, App</th>  | US-10-123-108-319  | Sequence 319, App |
|  | 31 | 1064   | 73.9  | 1266 | 15 <th>US-10-123-236-319</th> <th>Sequence 319, App</th>  | US-10-123-236-319  | Sequence 319, App |
|  | 32 | 1064   | 73.9  | 1266 | 15 <th>US-10-123-261-319</th> <th>Sequence 319, App</th>  | US-10-123-261-319  | Sequence 319, App |
|  | 33 | 1064   | 73.9  | 1266 | 15 <th>US-10-140-921-319</th> <th>Sequence 319, App</th>  | US-10-140-921-319  | Sequence 319, App |
|  | 34 | 1064   | 73.9  | 1266 | 15 <th>US-10-140-925-319</th> <th>Sequence 319, App</th>  | US-10-140-925-319  | Sequence 319, App |
|  | 35 | 1064   | 73.9  | 1266 | 15 <th>US-10-121-045-319</th> <th>Sequence 319, App</th>  | US-10-121-045-319  | Sequence 319, App |
|  | 36 | 1064   | 73.9  | 1266 | 15 <th>US-10-123-292-319</th> <th>Sequence 319, App</th>  | US-10-123-292-319  | Sequence 319, App |
|  | 37 | 1064   | 73.9  | 1266 | 15 <th>US-10-123-903-319</th> <th>Sequence 319, App</th>  | US-10-123-903-319  | Sequence 319, App |
|  | 38 | 1064   | 73.9  | 1266 | 15 <th>US-10-124-819-319</th> <th>Sequence 319, App</th>  | US-10-124-819-319  | Sequence 319, App |
|  | 39 | 1064   | 73.9  | 1266 | 15 <th>US-10-124-822-319</th> <th>Sequence 319, App</th>  | US-10-124-822-319  | Sequence 319, App |
|  | 40 | 1064   | 73.9  | 1266 | 15 <th>US-10-140-925-319</th> <th>Sequence 319, App</th>  | US-10-140-925-319  | Sequence 319, App |
|  | 41 | 1064   | 73.9  | 1266 | 15 <th>US-10-160-998-319</th> <th>Sequence 319, App</th>  | US-10-160-998-319  | Sequence 319, App |
|  | 42 | 1064   | 73.9  | 1266 | 15 <th>US-10-124-824-319</th> <th>Sequence 319, App</th>  | US-10-124-824-319  | Sequence 319, App |
|  | 43 | 1064   | 73.9  | 1266 | 15 <th>US-10-127-825-319</th> <th>Sequence 319, App</th>  | US-10-127-825-319  | Sequence 319, App |
|  | 44 | 1064   | 73.9  | 1266 | 15 <th>US-10-127-829-319</th> <th>Sequence 319, App</th>  | US-10-127-829-319  | Sequence 319, App |
|  | 45 | 1064   | 73.9  | 1266 | 15 <th>US-10-127-835A-319</th> <th>Sequence 319, App</th> | US-10-127-835A-319 | Sequence 319, App |

#### ALIGNMENTS

RESULT 1  
US-10-010-408-3  
Sequence 3, Application US/10010408  
Publication No. US20020165185A1  
GENERAL INFORMATION:  
APPLICANT: John J. Castelliott, Jr.  
TITLE OF INVENTION: NO. US20020165185A1 Heparin-Induced CCN-Like Molecules  
and Uses Therefor  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD, LLP  
STREET: 28 State Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/010,408  
FILING DATE: 07-Dec-2001  
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/044,273  
FILING DATE: March 19, 1998  
APPLICATION NUMBER: <Unknown>  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Amy E. Mandragouras  
REGISTRATION NUMBER: 36,207  
REFERENCE/DOCKET NUMBER: MBI-004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)742-4214  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 753 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..750  
SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
US-10-010-408-3  
Alignment Scores:  
Pred. No.: 4,26e-125 Length: 753  
Score: 1440.00 Matches: 250  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 0  
DB: 15  
US-10-010-408-2 (1-250) x US-10-010-408-3 (1-753)  
QY 1 MetArgGlySerProLeuIleHisLeuAlaThrSephLeuCysIleuSerMet 20  
DB 1 ATGAGGGGAGCCCGACCTGATCCATCTTCGGCCACTCTCTCCCTTCCTCAAG 60  
QY 21 ValCysAlaGlnLeuCysArgThrProCysThrCysProThrPthProProGlnCysPro 40  
DB 61 GTGTGCGCCACCTGTGGCGGACACCTGTACTGTCTCTGAGACACACCCAGAGGCCA 120  
QY 41 GlnGlyValProLeuValLeuAspGlyCysGlyCysValCysAlaArgArgLeu 60  
DB 121 CAGGGGTACCCCTGCTGTGATGCTGTGGCTGTCTGAAAGTGTGTGACGAGGCTG 180  
QY 61 GlyIleuSerCysAspHisLeuHisValCysAspProSephGlnGlyLeuValCysGlnPro 80  
DB 181 GGGGAGTCTGCGACCACTGCATGTCTGCGACCCCGAGCGAGGCTGTGTCAGCCT 240  
QY 81 GlyAlaGlyProGlnGlyHisGlyAlaValCysLeuLeuAspGlnAspAspGlySerCys 100  
DB 241 GGGGAGGCGCCGTGGGCGCATGGGCTGTGTCTCTTGATGAGATGACGATGACTGT 300  
QY 101 GluValAsnGlyArgArgTyrIleuAspGlyGluThrPheIysProAsnCysArgValLeu 120  
DB 301 GAGGTGAATGGCCGCGAGGTACTGATGATGAGACCTTTAAACCAATGACAGGGTCTCTG 360  
QY 121 CysArgCysAspAspGlyGlyPheThrCysLeuProLeuCysSerGlnAspValArgLeu 140  
DB 361 TGCCCTGTGATGAGGTGGCTTCACTGCTGCGCTGTGAGTGAGGATGTGGCGCTG 420  
QY 141 ProSerTrpAspCysProArgProLysArgIleGlnValProGlyLysCysCysProGln 160  
DB 421 CCCACTGGAGCTGCGCAAGCCCAAGAAATACAGGTCGCAAGAAAGTGTGCGCCGAG 480  
QY 161 TrpValCysAspGlnGlyValThrProAlaIleGlnArgSerThrAlaGlnGlyHisGln 180  
DB 481 TGGGTATGTGACCAAGGAGTACACCGCGGATCCAGCGCTCCACGCGCAAGACACCA 540  
QY 181 LeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThrAla 200

DB 541 CTTTGGCCCTGTACTCTCTGCTCTGCTGATAGCTCTTGTCCAAATTGGACACAGCC 600  
QY 201 TrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGlnAsn 220  
DB 601 TGGGGCCCTGTCTCAACACACCTGTGGCTGTGGCATAGCCACCGAGTGTCCAAACGAGAC 660  
QY 221 ArgPheCysGlnLeuGlnIleGlnArgArgLeuCysLeuProArgProCysLeuAlaAla 240  
DB 661 CGATTCTGCCAAGTGAATCCACAGCCCGCTGTCTGCCAGACCCCTGCGAGCC 720  
QY 241 ArgSerHisSerSerTrpAsnSerAlaPhe 250  
DB 721 AGGAGCCACAGCTCATGACAGACTGCTTTC 750  
RESULT 2  
US-10-010-408-1  
Sequence 1, Application US/10010408  
Publication No. US20020165185A1  
GENERAL INFORMATION:  
APPLICANT: John J. Castelli, Jr.  
TITLE OF INVENTION: No. US20020165185A1 Heparin-Induced CCN-Like Molecules  
and Uses Therefor  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD, LLP  
STREET: 28 State Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/010,408  
FILING DATE: 07-Dec-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/044,273  
FILING DATE: March 19, 1998  
APPLICATION NUMBER: <Unknown>  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Amy E. Mandragouras  
REGISTRATION NUMBER: 36,207  
REFERENCE/DOCKET NUMBER: MBI-004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)742-4214  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1708 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 249..1001  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-10-010-408-1  
Alignment Scores:  
Pred. No.: 1,05e-124 Length: 1708  
Score: 1440.00 Matches: 250  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 0  
DB: 15  
US-10-010-408-2 (1-250) x US-10-010-408-1 (1-1708)

QY 1 MetArglySerProLeuIleHisLeuLeuAlaThrSerPheLeuCysLeuLeuSerMet 20  
DB 249 ATGAGGGGAGAGCCACTGATCATCTTCGGCCACTTCCTTCCTGCTTCATG 308  
QY 21 ValCysAlaGlnLeuCysArgThrProCysThrCysProTrpThrProProGlnCysPro 40  
DB 309 GTGTGTGCCACTGCTGCGAGACCCCTGTACTCTGCTTGAGACACACCCAGTGCCTCA 368  
QY 41 GlnGlyValProLeuValLeuAspGlyCysGlyCysValCysAlaArgArgLeu 60  
DB 369 CAGGGGGTACCCCTGCTGTGATGATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 428  
QY 61 GlyIleSerCysAspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnPro 80  
DB 429 GGGGAGCTCTGGAGACACCTGATGTCTGCGACCCAGGAGGCTGCTGCTGCTGCTGCTGCT 488  
QY 81 GlyAlaGlyProGlyGlyHisGlyAlaValCysLeuLeuAspGlyAspAspGlySerCys 100  
DB 489 GGGGAGGCTCTGGAGACACCTGATGTCTGCGACCCAGGAGGCTGCTGCTGCTGCTGCTGCT 548  
QY 101 GluValAsnGlyArgArgThrLeuAspGlyGlyIleThrPheLeuProAsnCysArgValLeu 120  
DB 549 GAGGTGAATGGCCGAGGATGCTGATGAGAGACCTTAAACCAATTGACAGGCTCTG 608  
QY 121 CysArgCysAspAspGlyGlyPheThrCysLeuProLeuCysSerGlnAspValArgLeu 140  
DB 609 TCCCGCTGTGATGAGAGGTGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 668  
QY 141 ProSerTrpAspCysProArgProArgProArgGlnValProGlyLysCysCysProGln 160  
DB 669 CCCACTGGGACTGCCCCACGCCCCAAGAAATACAGGTCCAGGAAGTGGTCCGCGAG 728  
QY 161 TrpValCysAspGlnGlyValThrProAlaIleGlnArgSerThrAlaGlnGlyHisGln 180  
DB 729 TGGGTATGTGACAGGAGGATGACACCGGAGATCCAGCGCTCCACGCGCAAGAACCA 788  
QY 181 LeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThrAla 200  
DB 789 CTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 848  
QY 201 TrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGlnAsn 220  
DB 849 TGGGGCCCCCTGCTCAACACCTGCTGGGCTGAGATAGCACCCGAGTGTCCAAACAGAAC 908  
QY 221 ArgPheCysGlnLeuGlnIleGlnArgArgLeuCysLeuProArgProCysLeuAlaAla 240  
DB 909 CGATTCTGCAACTGAGATCCACGCGCCCTGTGTCTGCTGCCAGACCTGCTGCGAGCC 968  
QY 241 ArgSerHisSerSerTrpAsnSerAlaPhe 250  
DB 969 AGGAGCCACAGCTCATGAGACAGTCTTTC 998

RESULT 3  
US-10-010-408-12  
; Sequence 12, Application US/10010408  
; Publication No. US20020165185A1  
GENERAL INFORMATION:  
; APPLICANT: John J. Castelli, Jr.  
; TITLE OF INVENTION: No. US20020165185A1 Heparin-Induced CGN-like Molecules  
; and Uses Therefor  
NUMBER OF SEQUENCES: 13  
; CORRESPONDENCE ADDRESS:  
; ADDRESSER: LAHIVE & COCKFIELD, LLP  
; STREET: 28 State Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109  
COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/010,408  
; FILING DATE: 07-Dec-2001  
; CLASSIFICATION: <unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/044,273  
; FILING DATE: March 19, 1998  
; APPLICATION NUMBER: <unknown>  
; FILING DATE: <unknown>  
ATTORNEY/AGENT INFORMATION:  
; NAME: Amy E. Mandragouras  
; REGISTRATION NUMBER: 36,207  
; REFERENCE/DOCKET NUMBER: MBI-004  
TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617)227-7400  
; TELEFAX: (617)742-4214  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 681 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..681  
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:  
US-10-010-408-12  
Alignment Scores:  
Pred. No.: 2,83e-114 Length: 681  
Score: 1323.00 Matches: 227  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatch: 0  
Query Match: 91.88% Indels: 0  
DB: 15 Gaps: 0  
US-10-010-408-2 (1-250) x US-10-010-408-12 (1-681)  
QY 24 GlnLeuCysArgThrProCysThrCysProTrpThrProProGlnCysProGlnGlyVal 43  
DB 1 CAGGTGTCCGAGACACCTGATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 60  
QY 44 ProLeuValLeuAspGlyCysGlyCysValCysAlaArgArgLeuGlyGlySer 63  
DB 61 CCCGTGTGTGATGAGGT 120  
QY 64 CysAspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnProGlyValGly 83  
DB 121 TCGACACACCTGATGTCTGCGACCCAGCCAGGCGCTGTGTGTGTGTGTGTGTGTGTGTGT 180  
QY 84 ProGlyGlyHisGlyAlaValCysLeuLeuAspGlyAspAspGlySerCysGlnValAsn 103  
DB 181 CTTGCGCGCCATGAGGGCTGT 240  
QY 104 GlyArgArgThrLeuAspGlyGlyIleThrPheLeuProAsnCysArgValLeuCysArg 123  
DB 241 GCGCCGAGGTACTGATGAGAGAGACCTTAAACCAATTGAGGGCTCTGTGCGCGCTGT 300  
QY 124 AspAspGlyGlyPheThrCysLeuProLeuCysSerGlnAspValArgLeuProSerTrp 143  
DB 301 GATGACGGTGGCTTACCTGCTGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 360  
QY 144 AspCysProArgProGlyArgGlnGlnValProGlyLysCysCysProGlnTrpValCys 163  
DB 361 GACTGCCACCGCCCAAGAAATACAGGTGCGAGGAAAGTCTGCCCGAGTGGGTATGT 420  
QY 164 AspGlnGlyValThrProAlaIleGlnArgSerThrAlaGlnGlyHisGlnLeuSerAla 183  
DB 421 GACCAGGAGGTACACCGGAGATCCAGGCTTCCACGAGGAGACACCACTTCTCTGCC 480  
QY 184 LeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThrAlaTrpGlyPro 203

```

Db      481 CTTGTACTCTCTCTCTGTGATGCTCTTGTCCAAATTTGGAGCAGCAGCTGGGGCCC 540
QY      204 CysSerThrThrcysglyleuglylleaIaThrArvalSerAsnInglnaSerPheCys 223
Db      541 TGGTCAACCACTGTGGGCGTGGGCGATGACCCGAGTGTGCACCAACCAATCTCTGC 600
QY      224 GlnLeuGluIleGlnArvalGleuCysLeuProArpProCysLeuAlaIaArgSerHis 243
Db      601 CAACCTGAGATCCCAAGCCCGCTGTGTCTGCCAGACCCCTGTGGCAGCCAGAGCCAC 660
QY      244 SerSerTrpAsnSerAlaPhe 250
Db      661 AGCTCATGGAACAGTCTTTC 681

RESULT 4
US-10-112-267-17
: Sequence 17, Application US/10112267
: Publication No. US20030068678A1
: GENERAL INFORMATION:
: APPLICANT: Botstein, David A.
: APPLICANT: Cohen, Robert
: APPLICANT: Goddard, Audrey
: APPLICANT: Hillan, Kenneth J.
: APPLICANT: Levine, David A.
: APPLICANT: Lawrence, Diane
: APPLICANT: Pennica, Diane
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Wood, William I.
: TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
: FILE REFERENCE: P11762
: CURRENT APPLICATION NUMBER: US/10/112,267
: PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/182,145B
: PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-29
: PRIOR FILING DATE: EARLIER FILING DATE: 1997-10-29
: PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/063,704
: PRIOR FILING DATE: EARLIER FILING DATE: 1997-10-29
: PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/073,612
: PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-04
: PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/081,695
: PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-14
: NUMBER OF SEQ ID NOS: 156
: SEQ ID NO 17
: LENGTH: 1734
: TYPE: DNA
: ORGANISM: Mus musculus
US-10-112-267-17

Alignment Scores:
Pred. No.: 1,75e-112 Length: 1734
Score: 1308.50 Matches: 226
Percent Similarity: 93.63% Conservatve: 9
Best Local Similarity: 90.04% Mismatches: 15
Query Match: 90.87% Indels: 1
Gaps: 1

US-10-010-408-2 (1-250) x US-10-112-267-17 (1-1734)
QY      1 MetArgGlySerProleuIleHsLeuAlaIaThrSerPheLeuCysLeuSerMet 20
Db      257 ATGAGGGGCAACCACTGATCCATCTTGGCCATTCTCTCTGCAATCTCTCAATG 316
QY      21 ValCysAlaGlnLeuCysArgThrProCysThrCysProTrpThrProFrogInCysPro 40
Db      317 GTGTATTCCAGTGTGGCCAGACACCCCTGTGCTGTGAGACACACCCAGTGTGCA 376
QY      41 GlnGlyValProleuValLeuAspGlyCysGlyCysCysAlaValCysAlaArgArgLeu 60
Db      377 CCGGGGCTACCCCTGCTGTGATGCTGTGGCTGTGCTGTGCTGTGCTGTGCTGTG 436
QY      61 GlyCysSerCysAspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnPro 80

```

```

Db      437 GGGAGTCTCTGCGACCACTGATGCTGTGAGCCCGAGCCGCGCTGTGTCAGCT 496
QY      81 GlyAlaGlyProGlyLysGlyIleAlaValCysLeuLeuAspGlnAspArgCys 100
Db      497 GGGGAGAGCCCGCAGTGGCGGTGTGTGTGTCTTTTGGAGAGATGACGGAGCTGT 556
QY      101 GluValAsnGlyArgArgThrLeuAspGlyGluThrPheLysProAsnCysArgValLeu 120
Db      557 GAGTGAAATGGCCGCGAGTACCTGATGGGAGACCTTAAACCAATGACAGGGTTTG 616
QY      121 CysArgCysAspAspGlyGlyPheThrCysLeuProLeuCysSerTrpAsnValArgLeu 140
Db      617 TGGCTGTGTATGACCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 676
QY      141 ProSerTrpAspCysProArgProLysArgIleGlnValProGlyLysCysCysProGlu 160
Db      677 CCGAGCTGGGAGTCCCGCCAGCCCGCAGAGAAATACAGTGTCCAGAAAGTGTCTCC 736
QY      161 TrpValCysAspGlnGlyVal---ThrProAlaIleGlnArgSerThrAlaGlnGlyHis 179
Db      737 TGGGTGTGTAGACAGGAGTATGATGACCGCGCAATCCAGCCCTCTCAGCCCAAGAC 796
QY      180 GlnLeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThr 199
Db      797 CAACCTTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 856
QY      200 AlaTrpGlyProCysSerThrThrcysglyleuglylleaIaThrArvalSerAsnIngln 219
Db      857 GCTGTGGGCGCCCTGCTCAACCACTGTGGTGTGGATGACCCGAGATATTCACCA 916
QY      220 AsnArgPheCysGlnLeuGluIleGlnArvalGleuCysLeuProArpProCysLeuAla 239
Db      917 AACCGATTCGTGCAATGAGATCCAGCTGTGCTGTGTGTGTGTGTGTGTGTGTGT 976
QY      240 AlaArgSerHisSerSerTrpAsnSerAlaPhe 250
Db      977 TCAGAGGCGCAGCGCTCATGATGAAACAGTCTTTC 1009

RESULT 5
US-10-112-267-18/c
: Sequence 18, Application US/10112267
: Publication No. US20030068678A1
: GENERAL INFORMATION:
: APPLICANT: Botstein, David A.
: APPLICANT: Cohen, Robert
: APPLICANT: Goddard, Audrey
: APPLICANT: Hillan, Kenneth J.
: APPLICANT: Levine, David A.
: APPLICANT: Lawrence, Diane
: APPLICANT: Pennica, Diane
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Wood, William I.
: TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
: FILE REFERENCE: P11762
: CURRENT APPLICATION NUMBER: US/10/112,267
: PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/182,145B
: PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-29
: PRIOR FILING DATE: EARLIER FILING DATE: 1997-10-29
: PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/063,704
: PRIOR FILING DATE: EARLIER FILING DATE: 1997-10-29
: PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/073,612
: PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-04
: PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/081,695
: PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-14
: NUMBER OF SEQ ID NOS: 156
: SEQ ID NO 18
: LENGTH: 1734
: TYPE: DNA
: ORGANISM: Mus musculus
US-10-112-267-18

Alignment Scores:

```

Pred. No.: 1.75e-112 Length: 1734  
Score: 1308.50 Matches: 226  
Percent Similarity: 93.63% Conservatve: 9  
Best Local Similarity: 90.04% Mismatches: 15  
Query Match: 90.87% Indels: 1  
DB: 15 Gaps: 1

US-10-010-408-2 (1-250) x US-10-112-267-18 (1-1734)

QY 1 MetArgGlySerProLeuIleHisLeuAlaIrrSerPheLeuCysLeuLeuSerMet 20  
DB 1478 ATGAGGGGCAACCCACTGATCATCTTGGCCATTCTCTGATTCATCAATG 1419  
QY 21 ValCysAlaGlnLeuCysArgThrProCysThrCysProTrpPthProProGlnCysPro 40  
DB 1418 GGTGATTCACAGCTGTGCCACGACCCGTGCTCTCTTGACACCAACCCAGGCCCA 1359  
QY 41 GlnGlyValProLeuValLeuAspGlyCysGlyCysCysValCysAlaArgArgLeu 60  
DB 1358 CCGGGGGTACCCCTGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1299  
QY 61 GlyIleSerCysAspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnPro 80  
DB 1298 GGGGAGTCTTGGACACCTGATCTGTGCGACCCACCCAGGCGCTGTTGTAGCT 1239  
QY 81 GlyAlaGlyProGlyGlyHisGlyAlaValCysLeuLeuAspGlyAspAspGlySerCys 100  
DB 1238 GGGGAGGCCCCAGTGGCGGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1179  
QY 101 GluValAsnGlyArgArgTyrLeuAspGlyIuThrPheLysProAsnCysArgValLeu 120  
DB 1178 GAGGGAATGGCCCGAGGTACTGATGGGAGACCTTAAACCAATTGACAGGCTTTTG 1119  
QY 121 CysArgCysAspAspGlyGlyPheThrCysLeuProLeuAspGlySerGlnAspValArgLeu 140  
DB 1118 TGCCCTCTGATGAGGAGGTGTTTCACTGCTGCCGCTGCTGCTGCTGCTGCTGCTG 1059  
QY 141 ProSerTrpAspCysProArgProLysArgIleGlnValProGlyLysCysCysProGlu 160  
DB 1058 CCCACTGGAGCTGCCACGCCCCAGAGAAATACAGGTGCCAGAGAGGTGCTGCCCCAG 999  
QY 161 TrpValCysAspGlnGlyVal---ThrProAlaIleGlnArgSerThrAlaGlnGlyHis 179  
DB 998 TGGGTGTGTGACAGGACGATGTGACGCCGAATACCAACCCCTCTCCCAAGGAGACAC 939  
QY 180 GlnLeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThr 199  
DB 938 CACTTTCTGCCCTTGTCTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 879  
QY 200 AlaTrpGlyProCysSerThrThrcysGlyLeuGlyIleAlaThrArgValSerAsnGln 219  
DB 878 GCCTGGGGCCCTGCTCAACCAACCTGTGGTTGGGCAATGACCCCGATACCAACAG 819  
QY 220 AsnArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProArgProCysLeuAla 239  
DB 818 AACCATTTCTGCCAATGAGATCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 759  
QY 240 AlaArgSerHisSerSerTrpAsnSerAlaPhe 250  
DB 758 TCCAGAGCCACAGGCTCATGCAACAGTGCCTTC 726

RESULT 6  
US-10-137-866-319  
; Sequence 319, Application US/10137866  
; Publication No. US20030129689A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: Deforge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.

APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
TITLE OF INVENTION: ACIDS ENCODING THE SAME  
FILE REFERENCE: P3330R1C151  
CURRENT APPLICATION NUMBER: US/10/137,866  
PRIOR FILING DATE: 2002-05-03  
PRIOR APPLICATION NUMBER: 60/049911  
PRIOR FILING DATE: 1997-06-18  
PRIOR APPLICATION NUMBER: 60/056974  
PRIOR FILING DATE: 1997-08-26  
PRIOR APPLICATION NUMBER: 60/059113  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/059115  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/059117  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/059122  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/059184  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/059263  
PRIOR FILING DATE: 1997-09-18  
PRIOR APPLICATION NUMBER: 60/059352  
PRIOR FILING DATE: 1997-09-19  
PRIOR APPLICATION NUMBER: 60/059588  
PRIOR FILING DATE: 1997-09-19  
PRIOR APPLICATION NUMBER: 60/059836  
PRIOR FILING DATE: 1997-09-24  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/062285  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/062287  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/062814  
PRIOR FILING DATE: 1997-10-24  
PRIOR APPLICATION NUMBER: 60/062816  
PRIOR FILING DATE: 1997-10-24  
PRIOR APPLICATION NUMBER: 60/063045  
PRIOR FILING DATE: 1997-10-24  
PRIOR APPLICATION NUMBER: 60/063082  
PRIOR FILING DATE: 1997-10-31  
PRIOR APPLICATION NUMBER: 60/063127  
PRIOR FILING DATE: 1997-10-24  
PRIOR APPLICATION NUMBER: 60/063327  
PRIOR FILING DATE: 1997-10-27  
PRIOR APPLICATION NUMBER: 60/063329  
PRIOR FILING DATE: 1997-10-27  
PRIOR APPLICATION NUMBER: 60/063550  
PRIOR FILING DATE: 1997-10-28  
PRIOR APPLICATION NUMBER: 60/063561  
PRIOR FILING DATE: 1997-10-28  
PRIOR APPLICATION NUMBER: 60/063704  
PRIOR FILING DATE: 1997-10-29  
PRIOR APPLICATION NUMBER: 60/063733  
PRIOR FILING DATE: 1997-10-29  
PRIOR APPLICATION NUMBER: 60/063735  
PRIOR FILING DATE: 1997-10-29  
PRIOR APPLICATION NUMBER: 60/063738  
PRIOR FILING DATE: 1997-10-29  
PRIOR APPLICATION NUMBER: 60/063755  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/064248  
PRIOR FILING DATE: 1997-11-03





Db 250 GGGGACGAGACCCGGTGGCCGGGGGGCCCTGTGCTTGGCAGAGACGACGACGCTGT 309  
 QY 101 GluValAsnGlyArgTyrLeuAspGlyGluThrPheLysProAsnCysArgValLeu 120  
 Db 310 GAGGTGAACGGCCCGCTGATCGGGAGAGGAGACCTTCCAGCCCTGACATGACATCCG 369  
 QY 121 CysArgCysAspAspGlyGlyPheThrCysLeuProLeuCysSerGluAspValArgLeu 140  
 Db 370 TGGCGCTGCGAGGACGGCGGCTTACCTGCGCTGCGCTGTCAGCAGAGATGCGGCTG 429  
 QY 141 ProSerTrpAspCysProArgProLysArgGlyLeuValProGlyLysCysCysProGlu 160  
 Db 430 CCCAGCTGGAGCTCCCGCCACCCAGAGGCTCAGGCTCCGCGCAAGTGTGCTGCCAG 489  
 QY 161 TrpAlaCysAspGlnGlyValThrProAlaIleGlnArgSerThrAlaGlnGlyHisGln 180  
 Db 490 TGGGTGTGGCGCCAGAGAGGGGAGCTGGGAGACCCAGCCCTTCAGGCCCAAGAGCCAG 549  
 QY 181 LeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThrAla 200  
 Db 550 TTTTCTGGCCTTGTCTCTCCCTGCCCTGCTGTCCTGCCCTGCCAGATGGAGCAGGCC 609  
 QY 201 TrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGlnAsn 220  
 Db 610 TGGGAGCCCTGCTGACACCTGCGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 669  
 QY 221 ArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProArgProCysLeuAlaAla 240  
 Db 670 CGCTTGTGCGACCTGGAGACCCAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 729  
 QY 241 ArgSerHisSerSerTrpAsnSerAlaPhe 250  
 Db 730 AGGGGTGCGAGTCCACAAACAGTGCCTTC 759

RESULT 7

US-10-146-726-319  
 ; Sequence 319, Application US/10146726  
 ; Publication No. US20030129690A1  
 ; GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.  
 APPLICANT: Beresini, Maureen  
 APPLICANT: Deforge, Laura  
 APPLICANT: Desnoyers, Luc  
 APPLICANT: Filvaroff, Ellen  
 APPLICANT: Gao, Wei-Qiang  
 APPLICANT: Gerritsen, Mary E.  
 APPLICANT: Goddard, Audrey  
 APPLICANT: Godowski, Paul J.  
 APPLICANT: Gurney, Austin L.  
 APPLICANT: Sherwood, Steven  
 APPLICANT: Smith, Victoria  
 APPLICANT: Stewart, Timothy A.  
 APPLICANT: Tumas, Daniel  
 APPLICANT: Watanabe, Colin K  
 APPLICANT: Wood, William  
 APPLICANT: Zhang, Zemin  
 TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
 FILE REFERENCE: P3330R1C308  
 CURRENT APPLICATION NUMBER: US/10/146,726  
 PRIOR APPLICATION REMOVED - See File Wrapper or Palm  
 NUMBER OF SEQ ID NOS: 550  
 SEQ ID NO 319  
 LENGTH: 1266  
 TYPE: DNA  
 ORGANISM: Homo Sapien  
 US-10-146-726-319

Alignment Scores:  
 Pred. No.: 6.44e-90 Length: 1266  
 Score: 1064.00 Matches: 184  
 Percent Similarity: 80.40% Conservative: 17

Best Local Similarity: 73.60% Mismatches: 49  
 Query Match: 73.89% Indels: 0  
 DB: 14 Gaps: 0  
 US-10-010-408-2 (1-250) x US-10-146-726-319 (1-1266)

QY 1 MetArgLysSerProLeuIleHisLeuAlaAlaThrSerPheLeuCysLeuLeuSerMet 20  
 Db 10 ATGAGAGCAGACACCGAAGACCCAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 69  
 QY 21 ValCysAlaGlnLeuCysArgThrProCysThrCysProTrpPheProGlnCysPro 40  
 Db 70 GTGGCTACCCAGCTGTCGCCAGACCATGTGACCTGCCCCCGGCGCACCTCCCATGCCCC 129  
 QY 41 GlnGlyValProLeuValLeuAspGlyCysGlyCysCysValCysAlaArgArgLeu 60  
 Db 130 CTGGAGATACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 189  
 QY 61 GlyLeuSerCysAspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnPro 80  
 Db 190 GGGGAGCCCTGCGACCAACCTCCAGCTGGAGAGCCAGCCAGGGGCTGCTGCGACGCC 249  
 QY 81 GlyAlaGlyProGlyGlyHisGlyAlaValCysLeuLeuAspGlyAspAspGlySerCys 100  
 Db 250 GGGGACGAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 309  
 QY 101 GluValAsnGlyArgTyrLeuAspGlyGluThrPheLysProAsnCysArgValLeu 120  
 Db 310 GAGGTGAACGGCCCGCTGATCGGGAGAGGAGACCTTCCAGCCCTGACATGACATCCG 369  
 QY 121 CysArgCysAspAspGlyGlyPheThrCysLeuProLeuCysSerGluAspValArgLeu 140  
 Db 370 TGGCGCTGCGAGGACGGCGGCTTACCTGCGCTGCGCTGTCAGCAGAGATGCGGCTG 429  
 QY 141 ProSerTrpAspCysProArgProLysArgGlyLeuValProGlyLysCysCysProGlu 160  
 Db 430 CCCAGCTGGAGCTCCCGCCACCCAGAGGCTCAGGCTCCGCGCAAGTGTGCTGCCAG 489  
 QY 161 TrpAlaCysAspGlnGlyValThrProAlaIleGlnArgSerThrAlaGlnGlyHisGln 180  
 Db 490 TGGGTGTGGCGCCAGAGAGGGGAGCTGGGAGACCCAGCCCTTCAGGCCCAAGAGCCAG 549  
 QY 181 LeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThrAla 200  
 Db 550 TTTTCTGGCCTTGTCTCTCCCTGCCCTGCTGTCCTGCCCTGCCAGATGGAGCAGGCC 609  
 QY 201 TrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGlnAsn 220  
 Db 610 TGGGAGCCCTGCTGACACCTGCGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 669  
 QY 221 ArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProArgProCysLeuAlaAla 240  
 Db 670 CGCTTGTGCGACCTGGAGACCCAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 729  
 QY 241 ArgSerHisSerSerTrpAsnSerAlaPhe 250  
 Db 730 AGGGGTGCGAGTCCACAAACAGTGCCTTC 759

RESULT 8

US-10-146-727-319  
 ; Sequence 319, Application US/10146727  
 ; Publication No. US20030129691A1  
 ; GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.  
 APPLICANT: Beresini, Maureen  
 APPLICANT: Deforge, Laura  
 APPLICANT: Desnoyers, Luc  
 APPLICANT: Filvaroff, Ellen  
 APPLICANT: Gao, Wei-Qiang  
 APPLICANT: Gerritsen, Mary E.  
 APPLICANT: Goddard, Audrey  
 APPLICANT: Godowski, Paul J.  
 APPLICANT: Gurney, Austin L.



```

Qy 101 GluValaIaengIyArGArgTyrLeuAspGlyIuThrPhelYsProAsnCysArgValLeu 120
Db 310 GAGGTGAACGGCGCCCTGATCGGAGAGGGAGACCTTCAGCCCCACTGCAAGCTCCGC 369
Qy 121 CysArgCysAspAspGlyGlyPheThrCysLeuProLeuCysSerGluAspValArgLeu 140
Db 370 TGCCCTCCGAGAGAGGGCGCTTACCTGCGTGGCGCTGCGAGAGAGATGTGGCGCTG 429
Qy 141 ProSerTPAspCysProArgProlyArgIleGlnValProGlyLyCysCysProIu 160
Db 430 CCCAGCTGGAGACTGCCCCACCCAGAGGGGTCTGGGCAAGTGTGCTGCTAG 489
Qy 161 TrpValCysAspGlnGlyValThrProAlaIleGlnAspSerThrAlaGlnGlyHisGln 180
Db 490 TGGGTGTGCGGCAAGAGGGGAGTGGGAGCCGCCCTTCAGGCCCAAGACCCAG 549
Qy 181 LeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThrAla 200
Db 550 TTTTCTGGCTGTCTCTCTCCCTGCCCCCTGCTGCTGCCAGATGGAGCAGGCC 609
Qy 201 TrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGlnAsn 220
Db 610 TGGGACCCCTGCTGACCACTGTGGCTGGGCAATGGCACCCGGGTGTCCACAGAAC 669
Qy 221 ArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProArgProCysLeuAlaLa 240
Db 670 CGCTTCTGCGCACTGGAACCAAGCGCGCGCTGTGCTGCTCCAGAGCCCTGCCCCACCTCC 729
Qy 241 ArgSerHisSerSerTrpAsnSerAlaPhe 250
Db 730 AGGGGTCCGAGTCCACAAACAGTCCCTTC 759

```

## RESULT 10

```

US-10-152-380-319
; Sequence 319, Application US/10152380
; Publication No. US20030129694A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C397
; CURRENT APPLICATION NUMBER: US/10/152,380
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 319
; LENGTH: 1266
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-152-380-319

```

```

Alignment Scores:
Pred. No.: 6,44e-90
Score: 1064.00
Percent Similarity: 80.40%
Best Local Similarity: 73.60%
Length: 1266
Matches: 184
Conservative: 17
Mismatch: 49

```

```

Query Match: 73.89%
DB: 14
US-10-010-408-2 (1-250) x US-10-152-380-319 (1-1266)
Indels: 0
Gaps: 0

```

```

Qy 1 MetaArglySerProLeuIleHisLeuLeuAlaThrSerPheLeuCysLeuLeuSerMet 20
Db 10 ATGAGAGGACACACCGAAGACCCACTCTGCGCTTCTCCCTCTGCTGCTGCTGCTGCTG 69
Qy 21 ValCysAlaGlnLeuCysArgThrProCysThrCysProTrpThrProGlnCysPro 40
Db 70 GTGGCTACCCAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 129
Qy 41 GlnGlyValProLeuValLeuAspGlyCysGlyCysValCysValAlaArgArgLeu 60
Db 130 CTGGAGTACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 189
Qy 61 GlyIuSerCysAspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnPro 80
Db 190 GGGGAGCCCTGCGACCACTCCAGCTGCGAGCCAGCCAGGGGCTGCTGCTGCTGCTGCTG 249
Qy 81 GlyAlaGlyProGlyGlyHisGlyAlaValCysLeuLeuAspGluAspAspGlySerCys 100
Db 250 GGGGAGGACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 309
Qy 101 GluValaIaengIyArGArgTyrLeuAspGlyIuThrPhelYsProAsnCysArgValLeu 120
Db 310 GAGGTGAACGGCGCCCTGATCGGAGAGGGAGACCTTCAGCCCCACTGCAAGCTCCGC 369
Qy 121 CysArgCysAspAspGlyGlyPheThrCysLeuProLeuCysSerGluAspValArgLeu 140
Db 370 TGCCCTCCGAGAGAGGGCGCTTACCTGCGTGGCGCTGCGAGAGATGTGGCGCTG 429
Qy 141 ProSerTPAspCysProArgProlyArgIleGlnValProGlyLyCysCysProIu 160
Db 430 CCCAGCTGGAGACTGCCCCACCCAGAGGGGTCTGGGCAAGTGTGCTGCTAG 489
Qy 161 TrpValCysAspGlnGlyValThrProAlaIleGlnAspSerThrAlaGlnGlyHisGln 180
Db 490 TGGGTGTGCGGCAAGAGGGGAGTGGGAGCCGCCCTTCAGGCCCAAGACCCAG 549
Qy 181 LeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThrAla 200
Db 550 TTTTCTGGCTGTGCTCTCTCCCTGCCCCCTGCTGCTGCCAGATGGAGCAGGCC 609
Qy 201 TrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGlnAsn 220
Db 610 TGGGACCCCTGCTGACCACTGTGGCTGGGCAATGGCACCCGGGTGTCCACAGAAC 669
Qy 221 ArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProArgProCysLeuAlaLa 240
Db 670 CGCTTCTGCGCACTGGAACCAAGCGCGCGCTGTGCTGCTCCAGAGCCCTGCCCCACCTCC 729
Qy 241 ArgSerHisSerSerTrpAsnSerAlaPhe 250
Db 730 AGGGGTCCGAGTCCACAAACAGTCCCTTC 759

```

## RESULT 11

```

US-10-153-934-319
; Sequence 319, Application US/10153934
; Publication No. US20030129695A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven

```



;; PRIOR FILING DATE: 1997-10-27  
;; PRIOR APPLICATION NUMBER: 60/063329  
;; PRIOR FILING DATE: 1997-10-27  
;; PRIOR APPLICATION NUMBER: 60/063550  
;; PRIOR FILING DATE: 1997-10-28  
;; PRIOR APPLICATION NUMBER: 60/063561  
;; PRIOR FILING DATE: 1997-10-28  
;; PRIOR APPLICATION NUMBER: 60/063704  
;; PRIOR FILING DATE: 1997-10-29  
;; PRIOR APPLICATION NUMBER: 60/063733  
;; PRIOR FILING DATE: 1997-10-29  
;; PRIOR APPLICATION NUMBER: 60/063735  
;; PRIOR FILING DATE: 1997-10-29  
;; PRIOR APPLICATION NUMBER: 60/063738  
;; PRIOR FILING DATE: 1997-10-29  
;; PRIOR APPLICATION NUMBER: 60/063755  
;; PRIOR FILING DATE: 1997-10-17  
;; PRIOR APPLICATION NUMBER: 60/064248  
;; PRIOR FILING DATE: 1997-11-03  
;; PRIOR APPLICATION NUMBER: 60/064809  
;; PRIOR FILING DATE: 1997-11-07  
;; PRIOR APPLICATION NUMBER: 60/065186  
;; PRIOR FILING DATE: 1997-11-12  
;; PRIOR APPLICATION NUMBER: 60/065846  
;; PRIOR FILING DATE: 1997-11-17  
;; PRIOR APPLICATION NUMBER: 60/066364  
;; PRIOR FILING DATE: 1997-11-21  
;; PRIOR APPLICATION NUMBER: 60/066453  
;; PRIOR FILING DATE: 1997-11-24  
;; PRIOR APPLICATION NUMBER: 60/066511  
;; PRIOR FILING DATE: 1997-11-24  
;; PRIOR APPLICATION NUMBER: 60/066770  
;; PRIOR FILING DATE: 1997-11-24  
;; PRIOR APPLICATION NUMBER: 60/069212  
;; PRIOR FILING DATE: 1997-12-11  
;; PRIOR APPLICATION NUMBER: 60/069278  
;; PRIOR FILING DATE: 1997-12-11  
;; PRIOR APPLICATION NUMBER: 60/069334  
;; PRIOR FILING DATE: 1997-12-11  
;; PRIOR APPLICATION NUMBER: 60/069694  
;; PRIOR FILING DATE: 1997-12-16  
;; PRIOR APPLICATION NUMBER: 60/072320  
;; PRIOR FILING DATE: 1998-01-23  
;; PRIOR APPLICATION NUMBER: 60/073612  
;; PRIOR FILING DATE: 1998-02-04  
;; PRIOR APPLICATION NUMBER: 60/074086  
;; PRIOR FILING DATE: 1998-02-09  
;; PRIOR APPLICATION NUMBER: 60/074092  
;; PRIOR FILING DATE: 1998-02-09  
;; PRIOR APPLICATION NUMBER: 60/077791  
;; PRIOR FILING DATE: 1998-03-12  
;; PRIOR APPLICATION NUMBER: 60/078910  
;; PRIOR FILING DATE: 1998-03-20  
;; PRIOR APPLICATION NUMBER: 60/079294  
;; PRIOR FILING DATE: 1998-03-25  
;; PRIOR APPLICATION NUMBER: 60/079663  
;; PRIOR FILING DATE: 1998-02-27  
;; PRIOR APPLICATION NUMBER: 60/079728  
;; PRIOR FILING DATE: 1998-03-27  
;; PRIOR APPLICATION NUMBER: 60/080165  
;; PRIOR FILING DATE: 1998-03-31  
;; PRIOR APPLICATION NUMBER: 60/081203  
;; PRIOR FILING DATE: 1998-04-09  
;; PRIOR APPLICATION NUMBER: 60/081229  
;; PRIOR FILING DATE: 1998-04-09  
;; PRIOR APPLICATION NUMBER: 60/081695  
;; PRIOR FILING DATE: 1998-04-14  
;; PRIOR APPLICATION NUMBER: 60/081817  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/081818  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/082999  
;; PRIOR FILING DATE: 1998-04-24

;; PRIOR APPLICATION NUMBER: 60/083322  
;; PRIOR FILING DATE: 1998-04-28  
;; PRIOR APPLICATION NUMBER: 60/083545  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/084600  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084627  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084637  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/085149  
;; PRIOR FILING DATE: 1998-05-12  
;; PRIOR APPLICATION NUMBER: 60/085323  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085338  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085339  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085579  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085697  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085704  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/086414  
;; PRIOR FILING DATE: 1998-05-22  
;; PRIOR APPLICATION NUMBER: 60/086430  
;; PRIOR FILING DATE: 1998-05-22  
;; PRIOR APPLICATION NUMBER: 60/087106  
;; PRIOR FILING DATE: 1998-05-28  
;; PRIOR APPLICATION NUMBER: 60/088026  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088730  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088741  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088810  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088858  
;; PRIOR FILING DATE: 1998-06-11  
;; PRIOR APPLICATION NUMBER: 60/089532  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089599  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089907  
;; PRIOR FILING DATE: 1998-06-18  
;; PRIOR APPLICATION NUMBER: 60/089947  
;; PRIOR FILING DATE: 1998-06-19  
;; PRIOR APPLICATION NUMBER: 60/090349  
;; PRIOR FILING DATE: 1998-06-23  
;; PRIOR APPLICATION NUMBER: 60/090429  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090445  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090538  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090863  
;; PRIOR FILING DATE: 1998-06-26  
;; PRIOR APPLICATION NUMBER: 60/091360  
;; PRIOR FILING DATE: 1998-07-01  
;; PRIOR APPLICATION NUMBER: 60/091519  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091982  
;; PRIOR FILING DATE: 1998-07-07

## Alignment Scores:

|                        |          |               |      |
|------------------------|----------|---------------|------|
| Pred. No.:             | 6,44e-30 | Length:       | 1266 |
| Score:                 | 1064.00  | Matches:      | 184  |
| Percent Similarity:    | 80.40%   | Conservative: | 17   |
| Best Local Similarity: | 73.60%   | Mismatches:   | 49   |
| Query Match:           | 73.89%   | Indels:       | 0    |
| DB:                    | 15       | Gaps:         | 0    |

US-10-010-408-2 (1-250) x US-10-028-072-319 (1-1266)

QY 1 MetArgGlySerProLeuIleHisLeuAlaThrSerPheLeuCysLeuSerMet 20  
 Db ATGAGAGGACACCGAAGACCCACTCTCGGCTTCTCCCTCTCGGCTCTCAAG 69

QY 21 ValCysAlaGlnLeuCysArgThrProCysThrCysProTrpThrProGlnCysPro 40  
 Db GTGCGTACCCAGCTGTGCGGACACCATGTACTGCTCCCTGGCCACCTCCCGATGCCG 129

QY 41 GlnGlyValProLeuValLeuAspGlyCysGlyCysValCysAlaArgValLeu 60  
 Db CTGGGAGTACCCCTGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 189

QY 61 GlyGluSerCysAspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnPro 80  
 Db GGGGAGCCCTGGACACCACTCCACGCTCGAGAGCCAGGAGGCTGGTCTGCCAGCCC 249

QY 81 GlyAlaGlyProGlyGlyHisGlyAlaValCysLeuLeuAspGlnAspAspGlySerCys 100  
 Db GGGGAGGAGACCGGCTGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 309

QY 101 GluValAsnGlyArgArgThrLeuAspGlyGlyLeuThrPheLeuProAsnGlyValLeu 120  
 Db GAGGTACAGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 369

QY 121 CysArgCysAspAspGlyGlyPheThrCysLeuProLeuCysSerGlnAspValArgLeu 140  
 Db TCCCGCTCGAGAGACCGGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 429

QY 141 ProSerTrpAspCysProArgProGlyArgGlnValProGlyCysCysProGln 160  
 Db CCGAGCTGGAGACGCCCCACCCAGAGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 489

QY 161 TrpValCysAspGlnGlyValThrProAlaIleGlnArgSerThrAlaGlnHisGln 180  
 Db TGGGTGTGGGCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 549

QY 181 LeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThrAla 200  
 Db TTTTCTGGCTTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 609

QY 201 TrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgAlaSerAsnGlnAsn 220  
 Db TGGGAGCCCTGCTGCAACCACTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGT 669

QY 221 ArgPheCysGlnLeuGlnIleGlnArgArgLeuCysLeuProArgProCysLeuAlaAla 240  
 Db CGCTTCTGCGACTGTGAGACCCAGCGCGCTGTGCTGTCCAGAGCCCTGCCACCCCTCC 729

QY 241 ArgSerHisSerSerTrpAsnSerAlaPhe 250  
 Db AGGGGTGCGACGTCCACAAACAGTGCCTTC 759

Db 730 AGGGGTGCGACGTCCACAAACAGTGCCTTC 759

RESULT 13  
 US-10-121-049-319  
 ; Sequence 319, Application US/10121049  
 ; Publication No. US20030022239A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Baker, Kevin P.  
 ; APPLICANT: Beresini, Maureen  
 ; APPLICANT: Deforge, Laura  
 ; APPLICANT: Desnoyers, Luc  
 ; APPLICANT: Filvaroff, Ellen  
 ; APPLICANT: Gao, Wei-Qiang  
 ; APPLICANT: Geritsen, Mary E.  
 ; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Godowski, Paul J.  
 ; APPLICANT: Gurney, Austin L.  
 ; APPLICANT: Sherwood, Steven  
 ; APPLICANT: Smith, Victoria  
 ; APPLICANT: Stewart, Timothy A.  
 ; APPLICANT: Tumas, Daniel

APPLICANT: Watanabe, Colin K  
 APPLICANT: Wood, William  
 APPLICANT: Zhang, Zemin  
 TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
 TITLE OF INVENTION: ACIDS ENCODING THE SAME  
 FILE REFERENCE: P330R1C17  
 CURRENT APPLICATION NUMBER: US/10/121,049  
 PRIORITY FILING DATE: 2002-04-12  
 PRIORITY APPLICATION REMOVED - See File Wrapper or Palm  
 NUMBER OF SEQ ID NOS: 550  
 SEQ ID NO 319  
 LENGTH: 1266  
 TYPE: DNA  
 ORGANISM: Homo Sapien  
 US-10-121-049-319

Alignment Scores:  
 Pred. No.: 6,44e-90 Length: 1266  
 Score: 1064.00 Matches: 184  
 Percent Similarity: 80.40% Conservative: 17  
 Best Local Similarity: 73.60% Mismatches: 49  
 Query Match: 73.89% Indels: 0  
 DB: 15 Gaps: 0

US-10-010-408-2 (1-250) x US-10-121-049-319 (1-1266)

QY 1 MetArgGlySerProLeuIleHisLeuAlaThrSerPheLeuCysLeuSerMet 20  
 Db ATGAGAGGACACCGAAGACCCACTCTCGGCTTCTCCCTCTCGGCTCTCAAG 69

QY 21 ValCysAlaGlnLeuCysArgThrProCysThrCysProTrpThrProGlnCysPro 40  
 Db GTGCGTACCCAGCTGTGCGGACACCATGTACTGCTCCCTGGCCACCTCCCGATGCCG 129

QY 41 GlnGlyValProLeuValLeuAspGlyCysGlyCysValCysAlaArgValLeu 60  
 Db CTGGGAGTACCCCTGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 189

QY 61 GlyGluSerCysAspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnPro 80  
 Db GGGGAGCCCTGGACACCACTCCACGCTCGAGAGCCAGGAGGCTGGTCTGCCAGCCC 249

QY 81 GlyAlaGlyProGlyGlyHisGlyAlaValCysLeuLeuAspGlnAspAspGlySerCys 100  
 Db GGGGAGGAGACCGGCTGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 309

QY 101 GluValAsnGlyArgArgThrLeuAspGlyGlyLeuThrPheLeuProAsnGlyValLeu 120  
 Db GAGGTACAGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 369

QY 121 CysArgCysAspAspGlyGlyPheThrCysLeuProLeuCysSerGlnAspValArgLeu 140  
 Db TCCCGCTCGAGAGACCGGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 429

QY 141 ProSerTrpAspCysProArgProGlyArgGlnValProGlyCysCysProGln 160  
 Db CCGAGCTGGAGACGCCCCACCCAGAGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 489

QY 161 TrpValCysAspGlnGlyValThrProAlaIleGlnArgSerThrAlaGlnHisGln 180  
 Db TGGGTGTGGGCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 549

QY 181 LeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThrAla 200  
 Db TTTTCTGGCTTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 609

QY 201 TrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgAlaSerAsnGlnAsn 220  
 Db TGGGAGCCCTGCTGCAACCACTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGT 669

QY 221 ArgPheCysGlnLeuGlnIleGlnArgArgLeuCysLeuProArgProCysLeuAlaAla 240  
 Db CGCTTCTGCGACTGTGAGACCCAGCGCGCTGTGCTGTCCAGAGCCCTGCCACCCCTCC 729

```

Oy      241 ArgSerHisSerTrpAsnSerAlaPhe 250
        ||| | | | | | | | | | | | | | | |
Db      730 AGGGGTGCGAGTCCACAAAACAAGTGCTTC 759

RESULT 14
US-10-123-904-319
; Sequence 319, Application US/10123904
; Publication No. US20030022328A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gunney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P330RIC54
CURRENT APPLICATION NUMBER: US/10/123,904
PRIOR FILING DATE: 2002-04-16
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 319
LENGTH: 1266
TYPE: DNA
ORGANISM: Homo Sapien
US-10-123-904-319

Alignment Scores:
Pred. No.:          6.44e-90           Length:       1266
Score:              1064.00            Matches:       184
Percent Similarity: 80.40%             Conservative:  17
Best Local Similarity: 73.60%           Mismatches:   49
Query Match:         73.89%             Indels:        0
DB:                  15                 Gaps:          0
US-10-010-408-2 (1-250) x US-10-123-904-319 (1-1266)

Oy      1 MetArgLysEsrProLeuIleHIsLeuAlaThrSerPheLeuCysLeuLeuSerMet 20
        |||||:::|| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      10 ATGAGAGCAGCACCAAGAACACCACCTCCCTGCCTCTCCCTCTCTCTCTCAAG 69

Oy      21 ValCysAlaGlnCysArGthrProCysrThrCysProTrpThrProGlnCysPro 40
        || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      70 GTGGCTACCCAGCTGTGCCGACACCATATACCTGCCCTGGACCCTCCCGATGCCG 122

Oy      41 GlnGlyAlaProLeuValLeuAspGlyCysGlyCysLysValCysAlaArgArgLeu 60
        ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      130 CTGGGAATTACCCCTGGTGGCTGGATGGCTGTGCTCTGCGGGATGTGCAAGCGGCGTG 189

Oy      61 GlyLysEsrCysAspHisLeuHISValCysAspProSerGlnGlyLeuValCysGlnPro 80
        ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      190 GGGAGACCCCGACCACTCCACGTCCTCGACGCGCAGCAGCGAGCGCTGTGCGACGCC 249

Oy      81 GlyAlaGlyProoLygylHisGlyAlaValCysLeuLeuAspGluAspGlySerCys 100
        ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      250 GGCGCAGACCCCGCTGGCGCGGGGGGCGCTGTGCTCTTGCGAGAGACGACGACTGT 309

Oy      101 GluValasncGlyArgArgTrileuAspGlyGlnThrPheIysProAsnCysArgValLeu 122
        ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      310 GAAGTTAAACGGCGCGCTGTAATGCGGAAGGGGAGACCTTCCACCCCCTGAGCATCCGC 365

```

```

OY      12  CysAlcGysAspRspIgclylphheThrcySleuProLeucySSerGIuaSPvaIArGLeu 140
         |||||.....|||||.....|
Db      370  TGGCGTGCAGAGACGGCGCTTCACtCtGtGCCGTctCACGAGATGtCCGGCTG 4299
OY      141  ProSerTrpApScyProArGrProLyArGlIegInValProGlyLysCySCySProGlu 160
         |||||.....| | | | | .....| | | | |
Db      430  CCCAGCGTGAGACTGCCCCACCACCCAGAGAGGTGCAGAGTCTGGGCAAGTGTGCTGCCCTGAG 489
OY      161  TrpValcysAspGlnGlyValThrProAlaIIegInArGerThrAlaGlnGlyHisGln 180
         |||||.....| | | | | ..||| | | | | | | | | |
Db      490  TGGGTGTGCGCCCAAGAGAGGGGAGACTGTGGGAGACCCGCTTCACAGCCAAAGACCCCAG 549
OY      181  LeuSerAlaLeuValThrProAlaSerAlaAPraIaProcySPRoASnTrpSerThAla 200
         |||||.....| | | | | .....| | | | |
Db      550  TTTCCTGGCGCTGTCTCTTCCTCCCTGCCCCCTGTGTCCTGCTCCCAAGAATGGAGCGAGCC 6099
OY      201  TrpGlyProCySerThrThrcySGlyLeuGlyIleAlaThrArGvalSerAPsnGlnAsn 220
         |||||.....| | | | | .....| | | | |
Db      610  TGGGAGACCGTCTGCAGACACtGTGGGCTGGGCArTGGCACCCGGGTGTCCAACAGAAC 6659
OY      221  ArgPhcycSlmleuGluIegInArGaArgLeuCySleuProArGrProcySLeuaIAla 240
         |||||.....| | | | | .....| | | | | : : :
Db      670  CGCTTCGCGCATGTGAAGACCAAGCGCCGCTGTGCTGTCAAGGCCCTGCCACACCTCC 7299
OY      241  ArgSerHisSerTrpAsnSerAlaPhe 250
         ||| | | | | | | | | | | | | | |
Db      730  AGGGGTGCGAGTCCACAAMACAGTGCCTTC 759

RESULT 15
US-10-140-470-319
; Sequence 319, Application US/10140470
; Publication No. US2003002231A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Mei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OR INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: ACIDS ENCODING THE SAME
; CURRENT APPLICATION NUMBER: US/10/140,470
; CURRENT FILING DATE: 2002-05-06
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 319
; LENGTH: 1266
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-140-470-319

Alignment Scores:
Pred. No.:          6,44e-90           Length:       1266
Score:              1064.00            Matches:        184
Percent Similarity: 80.40%             Conservative:   17
Best Local Similarity: 73.60%           Mismatches:    49
Query Match:        73.89%             Indels:         0
DB:                  15                 Gaps:           0

US-10-010-408-2 (1-250) x US-10-140-470-319 (1-1266)
```

QY 1 MetArgGlySerProLeuIleHisLeuAlaThrSerPheLeuCysLeuLeuSerMet 20  
DB 10 ATGAGAGGACACACCGAAGACCCACTGCGCTCTCCGTCCTGCTCCCTCAAG 69  
QY 21 ValCysAlaGlnLeuysArgThrProCysThrCysProThrThrProProGlnCysPro 40  
DB 70 GTCCGTACCCAGCTGTGCCGACACCATCTACCTGCCCTGGCCACCTCCCCGATGCCG 129  
QY 41 GlnGlyValProLeuValLeuAspGlyCysGlyCysGlyValCysAlaArgArgLeu 60  
DB 130 CTGGGAGTACCCCTGTGCTGATGCTGTGGCTGTCCGCGGTATGTGCACGCGCTG 189  
QY 61 GlyGluSerCysAspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnPro 80  
DB 190 GGGGACCCCTGGCACCACCTCCACGCTCGACGACGACGAGGGCTGGTCTCCAGCCC 249  
QY 81 GlyAlaGlyProGlyIleHisGlyAlaValCysLeuLeuAspGlnAspAspGlySerCys 100  
DB 250 GGGGACGACCCCGGTGGCCGCGGCGCTCTGCTGCTGGCAGAGGACGACGACGCTGT 309  
QY 101 GluValAsnGlyArgArgThrLeuAspGlyGluThrPheGlyProAsnCysArgValLeu 120  
DB 310 GAGGTAAAGCGCCGCTGTATCGGGAAGGAGACCTTCGAGCCCACTGCACGATCCGC 369  
QY 121 CysArgCysAspAspGlyGlyPheThrCysLeuProLeuGlySerGlnAspValArgLeu 140  
DB 370 TCCCGCTCGAGAGACGGCGCTTCACCTCGTCCGCTGTCAGAGGATGTGCGCTG 429  
QY 141 ProSerTrpAspCysProArgProGlyArgIleGlnValProGlyIleCysCysProGlu 160  
DB 430 CCCAGCTGGGACTGCCCAACCCGACGAGGAGGTCGAGGCTGTGGCAAGTGTGCCCTGAG 489  
QY 161 TrpValCysAspGlnGlyValThrProAlaIleGlnArgSerThrAlaGlnGlyHisGln 180  
DB 490 TGGGTGTGGGCAAGAGGAGGAGGACTGGGAGACCCACCTTCACGACCCCAAGAGCCCA 549  
QY 181 LeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThrAla 200  
DB 550 TTTTCTGGCCTGTCTCTCCCTGCCCCCTGGTGTCCCTGCCCAAGATGAGACAGGCC 609  
QY 201 TrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGlnAsn 220  
DB 610 TGGGACCTGTGCTCGACCACTGTGGCTGTGGCATGCGCAACCGGGGTGTCCAACGAGAC 669  
QY 221 ArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProArgProCysLeuAlaAla 240  
DB 670 CGCTTCTGCGACTGAGAGACCGAGCGCCGCTGTGCTGTCCAGGCGCTGCCACCCCTCC 729  
QY 241 ArgSerHisSerSerTrpAsnSerAlaPhe 250  
DB 730 AGGGGTGCGAGTCCACAAAACAGTGCCTTC 759

Search completed: August 1, 2003, 04:51:29  
Job time : 237 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: August 1, 2003, 03:01:25 ; Search time 1455 Seconds  
(without alignments)  
2782.730 Million cell updates/sec

Title: US-10-010-408-2  
Perfect score: 1440  
Sequence: 1 MRGSPHLHLATSLFLCLISM.....LCPLRCLARSHSSWNSAF 250

Scoring table:  
BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 16154066 segs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+D2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO.spool/US10010408/runat\_25072003\_101856\_17934/app\_query.fasta\_1.391  
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blsum62 -TRANS=humana40.cdi -LIST=45  
-DOALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10010408.ecgn.1.1.1906.6runat.25072003\_101856\_17934 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEODUARY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6  
-Fgapext=7 -Ygapop=10 -Ygapext=0.5 -Delop=6 -Delext=7

Database :  
EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estlun:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hlc:\*  
9: gb\_estl:\*  
10: gb\_est2:\*  
11: gb\_hlc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estlom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vit:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rtd:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 1          | 1064  | 73.9        | 1058   | 14 | BM805088    |
| 2          | 1043  | 72.4        | 979    | 14 | BQ279131    |
| 3          | 977   | 67.8        | 886    | 13 | B1822142    |
| 4          | 927.5 | 64.4        | 750    | 13 | BM043988    |
| 5          | 918   | 63.7        | 1006   | 14 | BM921531    |
| 6          | 911   | 63.3        | 651    | 13 | B1457141    |
| 7          | 905   | 62.8        | 1073   | 14 | BQ073722    |
| 8          | 902.5 | 62.7        | 928    | 13 | B1161474    |
| 9          | 894   | 62.1        | 916    | 13 | B1457367    |
| 10         | 858   | 59.6        | 888    | 13 | B1825652    |
| 11         | 855.5 | 59.4        | 800    | 13 | B1826781    |
| 12         | 815   | 56.6        | 790    | 13 | BM046275    |
| 13         | 788   | 54.7        | 1022   | 14 | BQ952960    |
| 14         | 785   | 54.5        | 1166   | 13 | BG928868    |
| 15         | 751.5 | 52.2        | 1166   | 13 | BM543799    |
| 16         | 721   | 50.1        | 1251   | 14 | BM488499    |
| 17         | 721   | 50.1        | 1251   | 14 | BQ961357    |
| 18         | 696   | 48.3        | 749    | 9  | AL555144    |
| 19         | 694   | 48.2        | 452    | 10 | BE481184    |
| 20         | 668   | 46.4        | 933    | 14 | BQ278961    |
| 21         | 648   | 45.0        | 489    | 14 | BM751866    |
| 22         | 540.5 | 37.5        | 618    | 12 | BG538695    |
| 23         | 536   | 37.2        | 948    | 9  | AL542806    |
| 24         | 534   | 37.1        | 879    | 9  | AL547439    |
| 25         | 519   | 36.0        | 1002   | 9  | AL536435    |
| 26         | 498   | 34.6        | 1117   | 12 | BF982720    |
| 27         | 496.5 | 34.5        | 2029   | 11 | BC003205    |
| 28         | 490.5 | 34.1        | 753    | 12 | BE85071     |
| 29         | 486   | 33.8        | 1057   | 12 | BE898435    |
| 30         | 481.5 | 33.4        | 1898   | 11 | AK013300    |
| 31         | 473   | 32.8        | 940    | 14 | BQ937887    |
| 32         | 467   | 32.4        | 405    | 12 | BG900069    |
| 33         | 463.5 | 32.2        | 909    | 12 | BF981704    |
| 34         | 463.5 | 32.2        | 1076   | 13 | BM54624     |
| 35         | 460   | 31.9        | 936    | 14 | BQ880159    |
| 36         | 452.5 | 31.4        | 985    | 9  | AL550072    |
| 37         | 452   | 31.4        | 380    | 12 | BG900020    |
| 38         | 449   | 31.2        | 792    | 13 | B1823598    |
| 39         | 444.5 | 30.9        | 1085   | 13 | BM558306    |
| 40         | 444   | 30.8        | 693    | 13 | B1824200    |
| 41         | 444   | 30.8        | 973    | 13 | BM557871    |
| 42         | 442   | 30.7        | 826    | 12 | BG116760    |
| 43         | 441.5 | 30.7        | 836    | 14 | BQ716992    |
| 44         | 438.5 | 30.5        | 811    | 9  | AU133455    |
| 45         | 437   | 30.3        | 887    | 14 | BQ894536    |

## ALIGNMENTS

RESULT 1  
LOCUS BM805088 1058 bp MRNA linear EST 05-MAR-2002  
DEFINITION AGENCOURT 6490429 NIH\_MGC\_125 Homo sapiens cDNA clone IMAGE:5725937  
ACCESSION BM805088  
VERSION 5', mRNA sequence.  
KEYWORDS BM805088.1 GI:19121911  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1 (bases 1 to 1058)  
NIH-MGC <http://mgc.ncl.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL  
COMMENT

Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgaabs-remail.nih.gov  
Tissue Procurement: Invitrogen  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
plate: LLAM12717 row: j column: 18  
High quality sequence stop: 697.

**FEATURES**  
**source**

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5725937"
/clone_1id="NH_MGC_125"
/lab_host="DH10B"
/note="Organ: ovary (pool of 3); Vector: pCMV-Sport6;
Site_1: EcoRV (destroyed); Site_2: NotI; RNA source pool
of three ovaries, from females ranging in age from 38 to
49 yo. Library is oligo-dT primed and directionally cloned
(EcoRV site is destroyed upon cloning). Average insert
size 2.1 kb, insert size range 1-3.5 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 036."

```

**Alignment Scores:**

|                        |          |               |     |
|------------------------|----------|---------------|-----|
| Pred. NO.:             | 1,67e-77 | Length:       | 105 |
| Score:                 | 1064.00  | Matches:      | 184 |
| Percent Similarity:    | 80.408   | Conservative: | 17  |
| Best Local Similarity: | 73.608   | Mismatches:   | 49  |
| Query Match:           | 73.898   | Indels:       | 0   |
| DB:                    | 14       | Gaps:         | 0   |

US-10-010-408-2 (1-250) x BM805088 (1-1058)

1 MeFARG1SerProLeuIleHisLeuLeuAlaThrSerPheLeuCysLeuLeuSerMet 20  
 Db 9 ATGAGAGCGACACCGAAGACCACCTCTCGCGCTTCTCTCCCTCTCTCTCTCAAG 68  
 QY 21 ValCysAlaGlnLeuCysArgThrProCysThrCysProThrTrpProProGlnCysPro 40  
 Db 69 GTGCGTACCACGCTGTGCCCGACACCATGTACTCTCCCTGGCCACCTCCCGATGCCG 128  
 QY 41 GlnGlyValProLeuValLeuAspGlyCysGlyCysCysValCysAlaArgArgLeu 60  
 Db 129 CTGGGAGTACCCCTGTGCTGGATGGCTGTGCTCTCCCGGTATGTGCACGGCGGTG 188  
 QY 61 GlyIleSerCysAspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnPro 80  
 Db 189 GGGGAGACCCCTGGCAGCAACTCCACGTCGTGCAGCGACGACGAGGGCGCTGTGCACGCC 248  
 QY 81 GlyAlaGlyProGlyIleHisGlyAlaValCysLeuLeuAspGlnAspAspGlySerCys 100  
 Db 249 GGGGACGAGACCCGCTGTGACGGGGGCCCTGTGCTCTTGGCAGAGAGACGACGACTGT 308  
 QY 101 GlnValAsnGlyValArgTrpLeuAspGlyGlnThrPheLeuProAsnCysArgValLeu 120  
 Db 309 GAGCTGAACGGCGCGCTGTATCGGAGAGGGAGACCTTCCAGCCCACTGCAGCATCCG 368  
 QY 121 CysArgCysAspAspGlyGlyPheThrCysLeuProLeuCysSerGlnAspValArgLeu 140  
 Db 369 TGGCGGTGGAGAGACCGCGGCTTCACTCGCGTGCCTGTCAAGCAGAGATGTCCGGCTG 428  
 QY 141 ProSerTrpAspCysProArgProLeuArgIleGlnValProGlyIleCysCysProGlu 160  
 Db 429 CCACGCTGGAGACTCCCCACCCACCGAGAGGCTCGAGGCTCTGGGGAATGCTGCTCGTAG 488

|    |     |   |     |
|----|-----|---|-----|
| QY | 161 | TriPalYcAspArgInGlyValThrProAlaIleGlnArgSerThrAlaGlnIleYnGln  | 180 |
| Db | 489 | TGGGTGTGCGCCCAAGAGAGGGGAGCTGGGGACCCAGCCCTTCCAGCCCAAGAGCCCGAG  | 548 |
| QY | 181 | LeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThala   | 200 |
| Db | 549 | TTTTTGTGCGCTGTCTCCCTCCCTGCGCCCTGATGTCCTCCCTCCAGAAATGAGACAGGCC | 608 |
| QY | 201 | TrpGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGlnAsn  | 220 |
| Db | 609 | TGGGGACCCCTGCTGCACCACTGTGGCTGGGGCAATGGCCACCCGGGTGTCCAAACAGAAC | 668 |
| QY | 221 | ArgPheCysGlnLeuGlnIleGlnArgArgLeuCysLeuProArgProCysLeuAlaAla  | 240 |
| Db | 669 | CGCTTCTCCACACTGGAGAACCCAGCGCGCGCTGCTGCACGAGCCCGCCACACCTTC     | 728 |
| QY | 241 | ArgSerHisSerSerTrpAsnSerAla                                   | 250 |
| Db | 729 | AGGGGTGCGAGTCCACAAACAGTGCCTTC                                 | 758 |

|            |   |        |      |        |                 |
|------------|---|--------|------|--------|-----------------|
| RESULT 2   | BQ279131  | 979 bp | mRNA | linear | EST 07-MAY-2002 |
| LOCUS      | BQ279131  |        |      |        |                 |
| DEFINITION | BQ279131<br>AGNCCOURT_7046721 NIH_MGC_107 Homo sapiens CDNA clone IMAGE:5805819 |        |      |        |                 |
| ACCESSION  | BQ279131  |        |      |        |                 |
| VERSION    | BQ279131  |        |      |        |                 |
| KEYWORDS   | BQ279131.1 GI:20489339  |        |      |        |                 |
| SOURCE     | EST.  |        |      |        |                 |
| ORGANISM   | human.  |        |      |        |                 |
|            | Homo sapiens  |        |      |        |                 |

## FEATURES

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone IMAGE:5805819"
/clone_lib="NIH_MGC_107"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/name="Organ: breast; Vector: pORF7; Site_1: EcoRI;
Site_2: XhoI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: This is a NIH_MGC Library."

```

**Alignment Scores:**

|                        |          |               |     |
|------------------------|----------|---------------|-----|
| Pred. No.:             | 8.93e-76 | Length:       | 180 |
| Score:                 | 1043.00  | Matches:      | 16  |
| Percent Similarity:    | 80.66%   | Conservative: | 47  |
| Best Local Similarity: | 74.07%   | Mismatches:   | 0   |
| Query Match:           | 72.43%   | Indels:       | 0   |
| DB:                    | 14       | Gaps:         | 0   |



```

Db      480 GTGGCG-CAAGAGGGGGGACTGGGAGCCAGCCCTTCGAGCCCAAGACCCAGTTTC 538
Qy      162 ralaieualthrproalaSerAlaAspAlaProCysPysGlnSerThrAlaTrp1 202
Db      539 TGGCGTGTCTCTCCCTGCCCTGCTGTCCTCCCTCCAGATGAGAGCAGCGCGGG 598
Qy      202 yProCysSerThrThrCysGlyLeuGlyIleAlaThrValSerAsnGlnAsnArgph 222
Db      599 ACCCTGCTCGACCACTGTGGCTGGCGTGGCCACCCGCTGTCCACACCAAGACCGCTT 658
Qy      222 eCysLeuGluIleGlnArgLeuCysLeuProArgProCysLeuAlaAlaArgse 242
Db      659 CTGGCGAACTGGAGACACGAGCGCGCTGTGCTGTCCAGGCCCTGCCACCTCCAGGG 718
Qy      242 rHisSerSerTrpAsnSerAlaPhe 250
Db      719 TCGCAGTCCACAAACAGTGCCTTC 743

RESULT 4
BM043988
LOCUS      603620978F1 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:5446794 5',
DEFINITION mRNA sequence.
ACCESSION BM043988
VERSION    BM043988.1 GI:16773255
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 750)
AUTHORS   NIH-MGC http://mgi.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT    Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: DCTD/DTF
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LHC1929 row: k column: 19
            High quality sequence stop: 714.
            Location/Qualifiers
                1..750
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone_image="5446794"
                /clone_11b="NIH-MGC_40"
                /tissue_type="Carcinoma, cell line"
                /lab_host="DH10B (phage-resistant)"
                /note="Organ: prostate; Vector: pCMV7; Site_1: XhoI;
                Site_2: EcoRI; cDNA made by oligo-dT priming.
                Directionally cloned into EcoRI/XhoI sites using the
                following 5' adaptor: GGCACGAG(G). Library constructed by
                Ling Hong in the laboratory of Gerald M. Rubin (University
                of California, Berkeley) using ZAP-cDNA synthesis kit
                (Stratagene) and Superscript II RT (Life Technologies).
                Note: this is a NIH-MGC library."
BASE COUNT 112 a 267 c 246 g 125 t
ORIGIN

```

## Alignment Scores:

```

Pred. No.:      1,756-66      Length:      750
Score:          927.50      Matches:      173
Percent Similarity: 79.83%      Conservative: 17
Best Local Similarity: 72.69%      Mismatches: 46
Query Match:    64.41%      Indels:      5
DB:             13      Gaps:      0

```

US-10-010-408-2 (1-250) x BM043988 (1-750)

```

Qy      1 MetArgGlySerProLeuIleHisLeuLeuAlaTrpSerPheLeuCysLeuSerMet 20
Db      16 ATGAGAGGACACACCGAAGACCACTCCCTGGCTTCTCCCTCTGCTCTCTCAAG 75
Qy      21 ValCysAlaGlnLeuCysArgThrProCysThrCysProTrpThrProGlnCysPro 40
Db      76 GTGCGTACCCACCTTGTGCGCCGACCATGTACTCTCCCTCCGACCTCCCGAGGCGG 135
Qy      41 GlnGlyValProLeuValLeuAspGlyCysGlyCysValAlaCysAlaArgArgLeu 60
Db      136 CTGGAGATGCCCTGTGTGTGTATGATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 195
Qy      61 GlnGlySerCysAspHisLeuHisValCysAsp-ProSerGlnGlyLeuValCysGlnPr 80
Db      196 GGGAGAGCCCTGGACCACTCCACGCTCGACGACCATCCAGGAGCGCTGTGCTGCGAGCC 255
Qy      80 GlnGlyAlaGlyProGlnGlyHisGlyAlaValCysLeuLeuAspGlnAspArgGlySerCys 100
Db      256 CGGGGACAGACCCCGGTGGACGGGGGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 315
Qy      100 SGIuValAsnGlyArgArgGlyLeuAspGlyGlnThrPheCysProAsnCysArgValLe 120
Db      316 TGAGGTGAACGCCCGCTGTATCGGGAAGGGAGACCTTCCAGCCCACTGACGATCG 375
Qy      120 uCysArgCysAspAspGlyGlyPheThrCysLeuProLeuCysSerGlnAspValArgLe 140
Db      376 CTGGCGCTCGAGAGACGGGGGCTTCACCTGCGCGCTGTGCGAGAGATGTGCGGCT 435
Qy      140 uProSerTrpAspCysProArgProGlyArgGlnGlnValProGlyGlyCysCysProGln 160
Db      436 GCCCAGCTGGGAGCTCCCGCCACGAGGGGTTCAGAGCTCGGGCAAGTGTGCTGCGCA 495
Qy      160 uTrpValCysAspGlnGlyValThrProAlaIleGlnArgSerThr-AlaGlnGlyHisG 180
Db      496 GTGGGTGTGGCGCCACGAGAGGGGAGCTGGGAGCCCAAGCCCTTCCAGCCCAAGACCC 555
Qy      180 InLeuSerAlaLeuValThrProAlaSerAlaAspAlaProCysProAsnTrpSerThr 200
Db      556 ACTTTTCTGTGCTGTCTCTCCCTGCGCCCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 615
Qy      200 IaTrGlyProCysSerThrThrCysGlyLeuGlyIleAlaThrArgValSerAsnGln 220
Db      616 CTTGGGAGACCTGTGCGAC-ACCTGTGGCTGGGAGGAGCCCGGGGTGCTCAACAGCA 674
Qy      220 snArgPheCysGlnLeuGluIleGlnArgArgLeuCysLeuProArgPro 236
Db      675 ACCGTTCAGCCGACTGGAGAC-CAGCGCGCTG-TGCTGTGCCAGGCC 722

RESULT 5
BM921531
LOCUS      1006 bp mRNA linear EST 12-MAR-2002
DEFINITION AGENCOURT.6708025 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5753009
            5' mRNA sequence.
ACCESSION BM921531
VERSION    BM921531.1 GI:19371910
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 1006)
AUTHORS   NIH-MGC http://mgi.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT    Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:

```



```

QY 33 ProTPrThProProGInCysProGInGlyValProLeuValLeuAspGlyCysGlyCys 52
DB 66 CCTGCGCCACCTCCCGCCGCTGCGGAGTACCCCTGCTGCTGATGGCTGTGCTGCTC 125
QY 53 CysLysValCysAlaArgArgLeuGlyGlySerCysAspHisLeuHisValCysAspPro 72
DB 126 TCCCGGGATGTGCACAGCGGCTGGGGAGCCCTGCGACCACTCCACCTGCGCGAGCC 185
QY 73 SerGInGlyLeuValCysGInProGInGlyValGlyProGInGlyHisGlyAlaValCysLeu 92
DB 186 ACCCA-GGCGTGTGCTGCGAGCCCGGGGAGGAGACCCGGGAGCGGGGGCCCTGCTCCTC 244
QY 93 LeuAspGluAspAspGlySerCysGluValAsnGlyValArgTyrLeuAspGlyGluThr 112
DB 245 TTGGCAGAGAGACAGACAGCTGTGTGAGTGAACGGCGCGCTGTATCGGAGAGGAGAC 304
QY 113 PheLysProAsnCysArgValLeuGlyCysArgCysAspAspGlyGlyPheThrCysLeuPro 132
DB 305 TTCACGCCCCCTGCGACGATCCGCTGCTGCGAGAGCGGGGCTTCACTGCGCGCG 364
QY 133 LeuCysSerGluAspValArgLeuProSerTyrPaspCysProArgProLysArgIleGln 152
DB 365 CTGTGAGGAGGAGATGTGGGCTGCGCCACTGGAGACTGCCACCCAGGAGGAGCGAG 424
QY 153 ValProGlyLysCysCysProGluTyrValCysAspGlyGlyValThrProAlaIleGln 172
DB 423 GTCTGCGGAGATGTGCTGCTGAGTGGGTGCGCGCCGAGAGGAGGAGTGGGAGACCCAG 484
QY 173 ArgSerThrAlaGlnGlyHisGlnLeuSerAlaLeuValThrProAlaSerAlaAspAla 192
DB 485 CCCCTTCCAGCCCAAGAGACCCCACTTTTCTGCGCTTGTGCTTCCCTGCGCGCGTGT 544
QY 193 ProCysProAsnTyrSerThrAlaTyrPglProCysSerThrThrCysGlyLeuGlyIle 212
DB 545 CCTGCGCCAGATGTGAGCAGCGCTGGGAGACCTGCTGACCACTGTGGGTGGGCGAG 604
QY 213 AlaThrArgValSerAsnGlnAsnArgPheCysGlnLeuGlu 226
DB 605 GCCACCGGGGTGTCCAGCAACCGTTCGCGACGTGGAG 646

```

```

RESULT 7
BO073722 1073 bp mRNA linear EST 02-APR-2002
LOCUS AGENCOURT_7046577 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:5806602
DEFINITION 5', mRNA sequence.
ACCESSION BO073722.1 GI:19902768
VERSION EST.
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1073)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-riemail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM2047 row: k column: 19
High quality sequence stop: 517.
Location/Qualifiers
1..1073
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5806602"
/clone_1lb="NIH_MGC_101"

```

```

FEATURES
source
1..1073
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5806602"
/clone_1lb="NIH_MGC_101"

```

```

/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/Note="Organ: lung; Vector: pORF7; Site 1: EcoRI; Site 2:
XhoI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Library constructed by Ling Hong in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH-MGC Library."
BASE COUNT 166 a 384 c 347 g 176 t
ORIGIN
Alignment Scores:
Pred. No.: 1,91e-64 Length: 1073
Score: 905.00 Matches: 165
Percent Similarity: 77.358 Conservative: 16
Best Local Similarity: 70.518 Mismatches: 46
Query Match: 62.854 Indels: 7
DB: 14 Gaps: 2
US-10-010-408-2 (1-250) x BO073722 (1-1073)
QY 10 LeuAlaThrSerPheLeuGlySerMetValCysAlaGlnLeuCysArgThrPro 29
DB 2 CTGGGCTTCTCCCTCTGCTGCTCAAGGTCGACACCTGTGCGCCAGACCA 61
QY 30 CysThrCysProTyrThrProGInCysProGInGlyValProLeuValLeuAspGly 49
DB 62 TGTACCTGCCCTTGGCCACTCCCGATGCCGCGGGAGTACCCCTGCTGTGATGCC 121
QY 50 CysGlyCysCysValCysAlaArgArgLeuGlyGlySerCysAspHisLeuHisVal 69
DB 122 TGTGCTGCTGCGGGATGTGCACAGCGCGTGGGAGACCTGCGACCACTCCACGTC 181
QY 70 CysAspProSerGlnGlyLeuValCysGlnProGlyAlaGlyProGlyGlyHisGlyAla 89
DB 182 TCGGAGCGCCAGCGAGGCTGTGCGCGCCGGGAGAGACCGGTGGCGGGGGGCC 241
QY 90 ValCysLeuLeuAspGlyAspAspGlySerCysGluValAsnGlyValArgTyrLeuAsp 109
DB 242 CTGTGCTTCTTGGCAGAGAGACAGACAGCTGTGAGTGAACGCGCCCTGTATGGGGA 301
QY 110 GlyGluThrPheLysProAsnCysArgValLeuGlyCysArgCysAspAspGlyPheThr 129
DB 302 GGGGAGACCTTCCAGCCCACTGACAGATCCGCTGCGCGAGAGAGCGCGCTTCAAC 361
QY 130 CysLeuProLeuGlySerGluAspValArgLeuProSerTyrPaspCysProArgProLys 149
DB 362 TCGGTGCGCTGTGACGAGATGTGGCGCTGCCAGCTGGAGCTGCCCAACCCAGG 421
QY 150 ArgIleGlnValProGlyLysCysCysProGluTyrPvalCysAspGlnGlyValThrPro 169
DB 422 AGGTGCAAGCTTCCAGGCAAGTGTGCTGAGTGTGGGCGCAAGAGAGGAGACTG 481
QY 170 AlaIleGlnArgSerThrAlaGlnGlyHisGlnLeuSerAlaLeuValThrProAlaSer 189
DB 482 GGGACCAAGCCCTTCCAGCCCAAGACCCCACTTTTGTGCGCTGTCTTCCCTGCC 541
QY 190 AlaAspAlaProCysProAsnTyrSerThrAlaTyrPglProCysSerThrThrCysGly 209
DB 542 CCTGTGTCCCTTCCGCAAGAGAGAGCGGGCTGGGAGCCCTGTGCAACCACTGTGGG 601
QY 210 Leu-GlyIleAla---ThrArgValSerAsnGlnAsnArgPheCysGlnLeu-GluIleG 228
DB 602 CTGGGGCATGGCCACCGCGGTGTGCCACCCAGAAACCGCTTCTCCAGCTGGAGACCC 661
QY 228 InArgArg-LeuCysLeu-----ProArgPro 236
DB 662 AGCGCGCGCTGTGCTGAGCAGGACCGCTTGGCCCC 697
RESULT 8
B1161474

```

|  |  |               |      |        |                 |
|--|--|---------------|------|--------|-----------------|
| LOCUS                                      | BI161474   | 928 bp        | mRNA | linear | EST 05-JUN-2001 |
| DEFINITION                                 | 602864871p1 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:5019065 5', mRNA sequence.  |               |      |        |                 |
| ACCESSION                                  | BI161474   |               |      |        |                 |
| VERSION                                    | BI161474.1   | GI:14621475   |      |        |                 |
| KEYWORDS                                   | EST.   |               |      |        |                 |
| SOURCE                                     | human.   |               |      |        |                 |
| ORGANISM                                   | Homo sapiens   |               |      |        |                 |
| REFERENCE                                  | Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  |               |      |        |                 |
| AUTHORS                                    | NIH-MGC <a href="http://mgc.ncl.nih.gov/">http://mgc.ncl.nih.gov/</a> .  |               |      |        |                 |
| TITLE                                      | National Institutes of Health, Mammalian Gene Collection (MGC)   |               |      |        |                 |
| JOURNAL                                    | Unpublished (1999)   |               |      |        |                 |
| COMMENT                                    | Contact: Robert Strausberg, Ph.D.<br>Email: <a href="mailto:cgabs-remail.nih.gov">cgabs-remail.nih.gov</a><br>Tissue Procurement: ATCC<br>cDNA Library Preparation: Ling Hong/Rubin Laboratory<br>cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)<br>DNA Sequencing by: Incyte Genomics, Inc.<br>Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at:<br><a href="http://image.llnl.gov">http://image.llnl.gov</a><br>Plate: L16CM1833 row: m column: 18<br>High quality sequence start: 28<br>High quality sequence stop: 756.<br>Location/Qualifiers   |               |      |        |                 |
| FEATURES                                   | 1..928   |               |      |        |                 |
| SOURCE                                     | /organism="Homo sapiens"<br>/db_xref="taxon:9606"<br>/clone="IMAGE:5019065"<br>/clone_lib="NIH_MGC_42"<br>/tissue_type="epithelioid carcinoma cell line"<br>/lab_host="DH10B (phage-resistant)"<br>/note="Organ: pancreas; Vector: pOTB7; Site:1: XhoI; Site:2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH-MGC Library. 1" |               |      |        |                 |
| BASE COUNT                                 | 161 a 301 c 309 g 157 t  |               |      |        |                 |
| ORIGIN                                     |  |               |      |        |                 |
| Alignment Scores:                          |  |               |      |        |                 |
| Pred. No.:                                 | 2,56e-64   | Length:       | 928  |        |                 |
| Score:                                     | 902.50   | Matches:      | 172  |        |                 |
| Percent Similarity:                        | 77.14%   | Conservative: | 17   |        |                 |
| Best Local Similarity:                     | 70.20%   | Mismatches:   | 49   |        |                 |
| Query Match:                               | 62.67%   | Indels:       | 7    |        |                 |
| DB:  | 13   | Gaps:         | 1    |        |                 |
| 25-10-010-408-2 (1-250) x BI161474 (1-928) |  |               |      |        |                 |
| QY   | 8 HisleuLeuAlaThrSerPheIeuCysIleuLeuSerMetValCysAlaGlnIleuCysArg   | 27            |      |        |                 |
| DB   | 83 CACGTCGCGGCTTCCCTCCTCCTGCTCTCAAGATGTCGATACCCACTGTGGCCG  | 142           |      |        |                 |
| QY   | 28 ThrProCysThrCysProTyrThrProProGlnCysProGlnGlyValProIleuValIleu  | 47            |      |        |                 |
| DB   | 143 ACACCAATTAACCTGCGCCCTGGGACCACTCCCGATCCGCGGAGAGTACCCCTGTGCTG  | 202           |      |        |                 |
| QY   | 48 AspGlyCysGlyCysCysIleValCysAlaIleArgIleuGlyGlnSerCysAspHisIleu  | 67            |      |        |                 |
| DB   | 203 GATGGCTGTGCTGCTGCGGGGTATGTCACAGCGCGGCGGGAGACCCCTGCACCAACTGC  | 262           |      |        |                 |
| QY   | 68 HisValCysAspProSerGlnGlyIleuValCysGlnProGlyAlaGlyProGlyGlyHis   | 87            |      |        |                 |
| DB   | 263 CACGTCGCGACGCGACGAGCGCTGTGTCACAGCCCGGGGACAGACCCGCTGACGG  | 322           |      |        |                 |
| QY   | 88 GtAlaValCysLeuLeuAspGlnAspGlySerCysGluValAsnGlyArgArgTyr  | 107           |      |        |                 |

[illegible]







|           |   |
|-----------|---|
| KEYWORDS  | EST.  |
| SOURCE    | human.  |
| ORGANISM  | <i>Homo sapiens</i>   |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. |
| AUTHORS   | 1 (bases 1 to 790)  |
| TITLE     | NIH-MGC <a href="http://mgc.ncl.nih.gov/">http://mgc.ncl.nih.gov/</a> .   |
| JOURNAL   | National Institutes of Health, Mammalian Gene Collection (MGC)  |
| COMMENT   | Unpublished (1999)  |
|           | Contact: Robert Strausberg, Ph.D.   |

```

Plate: LLCM1944 row: h column: 17
High quality sequence stop: 780.
FEATURES
    Location/Qualifiers
        source          1..790

```

|            |       |       |       |       |
|------------|-------|-------|-------|-------|
| BASE COUNT | 121 a | 265 c | 269 g | 135 t |
| ORIGIN     |       |       |       |       |

```

Alignment Scores:
Pred. No.:      3.2e-57      Length:      790
Score:          815.00      Matches:     145
Percent Similarity: 78.82%   Conservative: 15
Best Local Similarity: 71.43% Mismatches:    40
Query Match:    56.60%      Indels:      4
DB:             13          Gaps:        1

US-10-010-408-2 (1-250) x BK046275 (1-790)

```



## COMMENT

Contact: Robert Strausberg, Ph.D.  
 Email: cgabbs-remail.nih.gov  
 Tissue Procurement: Invitrogen  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/BLNL at:  
 http://image.llnl.gov  
 Plate: LLM12361 row: f column: 15  
 High quality sequence stop: 547.

## FEATURES

Location/Qualifiers  
 1..1166  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5589134"  
 /clone\_1lb="NHL\_MGC\_125"  
 /lab\_host="DH10B"  
 /note="Organ: ovary (pool of 3); Vector: pCMV-Sport6;  
 Site\_1: EcoRV (destroyed); Site\_2: NotI; RNA source pool  
 of three ovaries, from females ranging in age from 38 to  
 49 yo. Library is oligo-dT primed and directionally cloned  
 (EcoRV site is destroyed upon cloning). Average insert  
 size 2.1 kb. Insert size range 1-3.5 kb. Library is  
 normalized and enriched for full-length clones and was  
 constructed by C. Gruber (Invitrogen). Research Genetics  
 tracking code 036."

BASE COUNT 216 a 424 c 315 g 207 t 4 others  
 ORIGIN

## Alignment Scores:

Pred. No.: 1.5e-54 Length: 1166  
 Score: 785.00 Matches: 149  
 Percent Similarity: 73.54% Conservative: 15  
 Best Local Similarity: 66.82% Mismatches: 43  
 Query Match: 54.51% Indels: 16  
 DB: 13 Gaps: 6

US-10-010-408-2 (1-250) x BM543799 (1-1166)

OY 1 MetArgGlySerProLeuIleHisLeuLeuAlaThrSerPheLeuCysLeuLeuSerMet 20  
 DB 211 ATGAGAGGACACACCGAAGACCACCTCTGCGCTTCTGCTCTGCTCTCTCAAG 270  
 OY 21 ValCysAlaGlnLeuCysArgThrProCysThrCysPheTyrPheProGlnCysPro 40  
 DB 271 GTGCGTACCGACGTGCGCGACACCATGTACCTGCCGCTGGCCACCTCCCGATGCCG 330  
 OY 41 GlnGlyValProLeuValLeuAspGlyCysGlyCysValCysAlaArgArgLeu 60  
 DB 331 CTGGAGTACCCCTGCTGCTGATGCTGTGCTGCTGCCGGTATGTGCAGCGCGCTG 390  
 OY 61 GlyIleSerCysAspHisLeuHisValCysAspProSerGlnGlyLeuValCysGlnPro 80  
 DB 391 GGGGAGCCCTGCGACCACTCCTGCGACGCCAGCCAGGCGCTGCTGCCAGCCC 450  
 OY 81 GlyAlaGlyProGlyGlyHisGlyAlaValCysLeuLeuAspGluAspAspGlySerCys 100  
 DB 451 GGGGAGAGCCCGGTGAGACGGGGGCGCTGCTGCTGCGCAGAGACAGACAGCTGT 510  
 OY 101 GluValAsnGlyArgArgTyrLeuAspGlyTyrPheLeysProAsnCysArgValLeu 120  
 DB 511 GAGGTGAACGGCCGCTGATCGGAGGAGGAGACCTTCCAGCCCACTGCAATCCGC 570  
 OY 121 CysArgCysAspAspGlyGlyPheThrCysLeuProLeuCysSerGluAspValArgLeu 140  
 DB 571 TGCCGCTCGAGAGAGCGGCGCTTACCTGCGCGCTGTGCGAGGAGAGATGTGCGCTG 630  
 OY 141 ProSerTyrAspCysProArg-ProLysArgIleGlnValProGlyLysCysCysProGln 160  
 DB 631 CCGAGCTGGAGTGGCCCGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 690  
 OY 160 uTrpValCysAspGln-----GlyValThrProAlaIleGlnArgSerThr 175

DB 691 ATGGGTGTGCGGCCAAAGAGGGGAGCTGGGGGACCAACCC-----TTCCA 738  
 OY 175 rAlaGlnGlyHisGln---LeuSerAlaLeuValThrPro---AlaSerAlaAspAlaPr 193  
 DB 739 ACCCAAGAGACCCCAAGNNTTTCTGTGGCTTGTCTCTTCCCGCGGCCCTGTGNGTCC 798  
 OY 193 oCys-ProAsnTrp---SerThrAlaTyrGlyProCysSerThrThrCys-----GlyL 210  
 DB 799 CTGGCCCAAAATGGAGAGCAACGGGCTGGGGAGACCTTGTCCAACAACCTGTGGGGCC 858  
 OY 210 euGly 211  
 DB 859 TGGGC 863

Search completed: August 1, 2003, 04:47:29  
 Job time : 1466 secs



New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3  
claim 33; Page 181-182; 284pp; English.

The present invention describes Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2 and WISP-3 have homology to connective tissue growth factor (CTGF). Products from the present invention can be used to treat WISP-related disorders such as breast, ovarian, and colon cancer or melanoma. The products can be used to treat arteriosclerosis. The products can also be used to treat other diseases e.g. benign and malignant tumours, leukaemia and lymphoid malignancies, neuronal, glial, astrocytal, hypothalamic and other glandular, macrophagal, epithelial, stromal, and blastocoealic disorders, haematopoiesis-related disorders, tissue-growth disorders, skin disorders, desmoplasia, fibrotic lesions, kidney disorders, bone-related disorders such as osteoporosis, trauma such as burns, incisions, and other wounds, connective tissue disorders, catabolic states, testicular-related disorders, and inflammatory, angiogenic and immunologic disorders including arteriosclerosis. The products can also be used for detection and diagnosis especially of individuals with neoplastic cell growth or proliferation. The products can be used in the production of transgenic or knock-out animals. Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing cells.

Sequence 251 AA;

Query Match 90.9%; Score 1308.5; DB 20; Length 251;  
Best Local Similarity 90.0%; Pred. No. 1e-90;  
Matches 226; Conservative 9; Mismatches 15; Indels 1; Gaps 1;

QY 1 MRGSPILHLATSEFLCLSMVCAQLCRPTCTPMPPOGVPVLVDGCGCCCKVCARL 60  
DB 1 MRGNPLHLALISFLCLISLMVYSQLCPACPMPTPPGPGVPLVLDGCGCCCKVCARL 60  
QY 61 GESCDHLHVCDPSQGLVCPGAGPGGAGVCLDEDDGSCVNGRRYLDGETFKRNCRYL 120  
DB 61 GESCDHLHVCDPSQGLVCPGAGPGGAGVCLDEDDGSCVNGRRYLDGETFKRNCRYL 120  
QY 121 CRDDGGFTCLPLCSEEDVRLPSPMDPRPRRIQVPGCCPEWVCDVAWMPALQPSAAGH 179  
DB 121 CRDDGGFTCLPLCSEEDVRLPSPMDPRPRRIQVPGCCPEWVCDVAWMPALQPSAAGH 180  
QY 180 QLSALVTPASADAPCPNMTAMGPGSTTCGIGIATRVSNQNRFCOLEIQRRLCLRPCL 239  
DB 181 QLSALVTPASADGCPNMTAMGPGSTTCGIGIATRVSNQNRFCOLEIQRRLCLRPCL 240  
QY 240 ARSHSSWNSAF 250  
DB 241 SRSHGWSNSAF 251

RESULT 4

AA17680  
ID AA17680 standard; Protein: 250 AA.

AA17680:

06-AUG-1999 (first entry)

Mouse WISP-2 protein SEQ ID NO:78.

WN-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;  
connective tissue growth factor; cancer; melanoma; arteriosclerosis;  
leukaemia; lymphoid malignancy; haematopoiesis-related disorder;  
tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;  
kidney disorder; bone-related disorder; osteoporosis; trauma; burn;  
connective tissue disorder; catabolic state; inflammation;  
testicular-related disorder; angiogenesis; immunological disorder.

Mus sp.  
OS  
XX

PN W09921998-A1.

XX 06-MAY-1999.

XX 29-OCT-1998; 98WO-US22991.

XX 14-APR-1998; 98US-0081695.

XX 29-OCT-1997; 97US-0063704.

XX 03-FEB-1998; 98US-0073612.

XX (GETH ) GENENTECH INC.

XX Botstein DA, Cohen RL, Goddard A, Gurney AL, Hillan K;  
XX Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WI;

XX WPI, 1999-337420/28.

XX New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3

XX Disclosure: Page 245-246; 284pp; English.

The present invention describes Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2 and WISP-3 have homology to connective tissue growth factor (CTGF). Products from the present invention can be used to treat WISP-related disorders such as breast, ovarian, and colon cancer or melanoma. The products can be used to treat arteriosclerosis. The products can also be used to treat other diseases e.g. benign and malignant tumours, leukaemia and lymphoid malignancies, neuronal, glial, astrocytal, hypothalamic and other glandular, macrophagal, epithelial, stromal, and blastocoealic disorders, haematopoiesis-related disorders, tissue-growth disorders, skin disorders, desmoplasia, fibrotic lesions, kidney disorders, bone-related disorders such as osteoporosis, trauma such as burns, incisions, and other wounds, connective tissue disorders, catabolic states, testicular-related disorders, and inflammatory, angiogenic and immunologic disorders including arteriosclerosis. The products can also be used for detection and diagnosis especially of individuals with neoplastic cell growth or proliferation. The products can be used in the production of transgenic or knock-out animals. Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing cells.

Sequence 250 AA;

Query Match 90.5%; Score 1303.5; DB 20; Length 250;  
Best Local Similarity 90.0%; Pred. No. 2.5e-90;  
Matches 225; Conservative 9; Mismatches 15; Indels 1; Gaps 1;

QY 2 RGSPLIHLATSEFLCLSMVCAQLCRPTCTPMPPOGVPVLVDGCGCCCKVCARL 61  
DB 1 RGNPLHLALISFLCLISLMVYSQLCPACPMPTPPGPGVPLVLDGCGCCCKVCARL 60  
QY 62 ESCDHLHVCDPSQGLVCPGAGPGGAGVCLDEDDGSCVNGRRYLDGETFKRNCRYL 121  
DB 61 ESCDHLHVCDPSQGLVCPGAGPGGAGVCLDEDDGSCVNGRRYLDGETFKRNCRYL 120  
QY 122 RCDGGFTCLPLCSEEDVRLPSPMDPRPRRIQVPGCCPEWVCDVAWMPALQPSAAGH 180  
DB 121 RCDGGFTCLPLCSEEDVRLPSPMDPRPRRIQVPGCCPEWVCDVAWMPALQPSAAGH 180  
QY 181 LSAVTPASADAPCPNMTAMGPGSTTCGIGIATRVSNQNRFCOLEIQRRLCLRPCL 240  
DB 181 LSAVTPASADGCPNMTAMGPGSTTCGIGIATRVSNQNRFCOLEIQRRLCLRPCL 240  
QY 241 RSHSSWNSAF 250  
DB 241 RSHGWSNSAF 250

RESULT 5

AA17681  
ID AA17681 standard; Protein: 249 AA.

XX

XX Claim 24; Fig 1; 108bp; English.  
 PS  
 CC This represents a rat heparin-induced CGN-like protein (HICP) protein.  
 CC Agents that stimulate or inhibit HICP protein activity or expression,  
 CC antisense HICP nucleic acid molecules and HICP antibodies, can be used to  
 CC modulate cell-associated activity. HICP modulators can be used to treat  
 CC disorders characterized by aberrant HICP protein activity or expression.  
 CC Probes capable of hybridizing to HICP mRNA or antibodies specific for  
 CC HICP can be used to detect HICP activity in a biological sample. HICP  
 CC can be used to treat disorders, such as a cardiovascular or fibrotic  
 CC disorder, characterized by aberrant cell proliferation.  
 XX  
 SQ Sequence 250 AA;

Query Match 100.0%; Score 1440; DB 20; Length 250;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-100;  
 Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MGSPLHLATSFCLLSMVCACQLCRPTCTCPMTTPPCPGVPLVDGCGCKVCARL 60  
 DB 1 MGSPLHLATSFCLLSMVCACQLCRPTCTCPMTTPPCPGVPLVDGCGCKVCARL 60  
 OY 61 GSCCHLVCDPSQGLVCPGAGPGSHGAVCLLDEDDGSCENVGRYLDGETFKPCRYL 120  
 DB 61 GSCCHLVCDPSQGLVCPGAGPGSHGAVCLLDEDDGSCENVGRYLDGETFKPCRYL 120  
 OY 121 CCGDGGFTCLPLCSEDEVRLPSMDCPRPRKRIQVPGKCEPWYCDGVTPAIORSTAGHQ 180  
 DB 121 CCGDGGFTCLPLCSEDEVRLPSMDCPRPRKRIQVPGKCEPWYCDGVTPAIORSTAGHQ 180  
 OY 181 LSALVTPASADAPCPNMSTAMGPCSTTCGLGIATRVSNQNRFCOLEIORLCLPRCLAA 240  
 DB 181 LSALVTPASADAPCPNMSTAMGPCSTTCGLGIATRVSNQNRFCOLEIORLCLPRCLAA 240  
 OY 241 RSHSSMNSAF 250  
 DB 241 RSHSSMNSAF 250

RESULT 2  
 ID AAY27440 standard; Protein; 227 AA.  
 AC AAY27440;

XX 26-NOV-1999 (first entry)  
 DE Rat HICP mature polypeptide.  
 XX Heparin-induced CGN-like protein; HICP; cell-associated activity;  
 KW cardiovascular disorder; aberrant cell proliferation; fibrotic disorder.  
 XX Rattus sp.  
 OS Mus sp.  
 XX WO9947556-A2.  
 XX 23-SEP-1999.  
 PD 18-MAR-1999; 99MO-US05999.  
 XX 19-MAR-1998; 98US-0044273.  
 XX (TUFT ) TUFTS COLLEGE.  
 XX Castellot JJ;

PS Disclosure; Fig 2; 108bp; English.  
 XX  
 CC The invention provides a rat heparin-induced CGN-like protein (HICP)  
 CC protein. Agents that stimulate or inhibit HICP protein activity or  
 CC expression, antisense HICP nucleic acid molecules and HICP antibodies,  
 CC can be used to modulate cell-associated activity. HICP modulators can be  
 CC used to treat disorders characterized by aberrant HICP protein activity  
 CC or expression. Probes capable of hybridizing to HICP mRNA or antibodies  
 CC specific for HICP can be used to detect HICP activity in a biological  
 CC sample. HICP can be used to treat disorders, such as a cardiovascular or  
 CC fibrotic disorder, characterized by aberrant cell proliferation. The  
 CC present sequence represents the rat HICP mature polypeptide.  
 XX  
 SQ Sequence 227 AA;

Query Match 91.9%; Score 1323; DB 20; Length 227;  
 Best Local Similarity 100.0%; Pred. No. 7.7e-92;  
 Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 24 QLCRTPTCTPMTTPPCPGVPLVDGCGCKVCARLSECDHLHVCDPQGLVCPGAG 83  
 DB 1 QLCRTPTCTPMTTPPCPGVPLVDGCGCKVCARLSECDHLHVCDPQGLVCPGAG 60  
 OY 84 PGHGAVALLEDDGSCENVGRYLDGETFKPCRYLRCDDGFTCLPLCSEDEVRLPSW 143  
 DB 61 PGHGAVALLEDDGSCENVGRYLDGETFKPCRYLRCDDGFTCLPLCSEDEVRLPSW 120  
 OY 144 DCPRRKRIQVPGKCEPWYCDGVTPAIORSTAGHQSALVTPASADAPCPNMSTAMGP 203  
 DB 121 DCPRRKRIQVPGKCEPWYCDGVTPAIORSTAGHQSALVTPASADAPCPNMSTAMGP 180  
 OY 204 CSTTCGLGIATRVSNQNRFCOLEIORLCLPRCLAAARSHSSMNSAF 250  
 DB 181 CSTTCGLGIATRVSNQNRFCOLEIORLCLPRCLAAARSHSSMNSAF 227

RESULT 3  
 ID AAY17651 standard; Protein; 251 AA.  
 AC AAY17651;

XX 06-AUG-1999 (first entry)  
 DE Mouse WISP-2 protein SEQ ID NO:20.  
 XX  
 KW WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;  
 KW connective tissue growth factor; cancer; melanoma; arteriosclerosis;  
 KW leukaemia; lymphoid malignancy; haematopoiesis-related disorder;  
 KW tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;  
 KW kidney disorder; bone-related disorder; osteoporosis; trauma; burn;  
 KW connective tissue disorder; catabolic state; inflammation;  
 KW testicular-related disorder; angiogenesis; immunological disorder.  
 XX  
 OS Mus sp.  
 XX WO9921998-A1.  
 XX 06-MAY-1999.  
 PD 29-OCT-1998; 98MO-US22991.  
 XX 14-APR-1998; 98US-0081695.  
 PR 29-OCT-1997; 97US-0063704.  
 PR 03-FEB-1998; 98US-0073612.  
 XX (GETH ) GENENTECH INC.  
 XX Botstein DA, Cohen RL, Goddard A, Gurney AL, Hillan K;  
 PI Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WI;  
 XX WPI: 1999-337420/28.  
 DR N-PSDB; AAX76488.

[illegible]



XX Sequence 248 AA;  
 SQ Query Match 89.8%; Score 1292.5; DB 20; Length 248;  
 Best Local Similarity 89.9%; Pred. No. 1.6e-89;  
 Matches 223; Conservative 9; Mismatches 15; Indels 1; Gaps 1;

QY 4 SPLHLATSFCLLSMCAQLCRPTCTCPMTPOCPGCVPLVLDGCGCCCKVCARRLGEES 63  
 DB 1 NPLHLATSFCLLSMYSQLCRPAACPMTPPOCPGCVPLVLDGCGCCCKVCARRLGEES 60  
 QY 64 CDHLHVCDSGGLVCPGAGPGGHCAGVCLLDEDDSCCEVNGRRYLDGETFKNCRVLCRC 123  
 DB 61 CDHLHVCDSGGLVCPGAGPGGHCAGVCLLDEDDSCCEVNGRRYLDGETFKNCRVLCRC 120  
 QY 124 DDGFTCLPLCSEVDRLPSMDCPRRRIOVPGRCCPEWVCDQAVMPAIOPSAOGHQLS 182  
 DB 121 DDGFTCLPLCSEVDRLPSMDCPRRRIOVPGRCCPEWVCDQAVMPAIOPSAOGHQLS 180  
 QY 183 ALVTPASADAPCPMNSTAMGPGSTTCGLGIATRVSNONRFOCLEIQRRLCLSRCLASRS 242  
 DB 181 ALVTPASADAPCPMNSTAMGPGSTTCGLGIATRVSNONRFOCLEIQRRLCLSRCLASRS 240  
 QY 243 HSNWSAF 250  
 DB 241 HSNWSAF 248

RESULT 7  
 AAY17683  
 ID AAY17683 standard; Protein: 247 AA.  
 AC AAY17683;  
 XX 06-AUG-1999 (first entry)  
 DT Mouse WISP-2 protein SEQ ID NO:81.  
 XX  
 KW Wnt-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;  
 KW connective tissue growth factor; cancer; melanoma; arteriosclerosis;  
 KW leukaemia; lymphoid malignancy; haematopoiesis-related disorder;  
 KW tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;  
 KW kidney disorder; bone-related disorder; osteoporosis; trauma; burn;  
 KW connective tissue disorder; catabolic state; inflammation;  
 KW testicular-related disorder; angiogenesis; immunological disorder.  
 XX  
 OS Mus sp.  
 XX  
 PN WO9921998-A1.  
 XX  
 PD 06-MAY-1999.  
 DT 29-OCT-1998; 98WO-US22991.  
 XX  
 PF 14-APR-1998; 98US-0081695.  
 PR 29-OCT-1997; 97US-0063704.  
 PR 03-FEB-1998; 98US-0073612.  
 XX  
 PA (GETH) GENENTECH INC.  
 XX  
 PI Botstein DA, Cohen RL, Goddard A, Gurney AL, Hillan K;  
 PI Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WI;  
 XX  
 DR WPI: 1999-337420/28.  
 XX  
 XX New isolated wnt-1 induced secreted polypeptides, WISP-1, 2 and 3  
 PT  
 PS Disclosure: Page 248; 284pp; English.  
 XX  
 CC The present invention describes wnt-1 induced secreted polypeptides,  
 CC WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2  
 CC and WISP-3 have homology to connective tissue growth factor (CTGF).  
 CC Products from the present invention can be used to treat WISP-related

CC disorders such as breast, ovarian, and colon cancer or melanoma. The  
 CC products can be used to treat arteriosclerosis. The products can also be  
 CC used to treat other diseases e.g. benign and malignant tumours,  
 CC leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,  
 CC hypothalamic and other glandular, macrophagal, epithelial, stromal, and  
 CC blastocoealic disorders, haematopoiesis-related disorders, tissue-growth  
 CC disorders, skin disorders, desmoplasia, fibrotic lesions, kidney  
 CC disorders, bone-related disorders such as osteoporosis, trauma such as  
 CC burns, incisions, and other wounds, connective tissue disorders,  
 CC catabolic states, testicular-related disorders, and inflammatory,  
 CC angiogenic and immunologic disorders including arteriosclerosis. The  
 CC products can also be used for detection and diagnosis especially of  
 CC individuals with neoplastic cell growth or proliferation. The products  
 CC can be used in the production of transgenic or knock-out animals.  
 CC Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing  
 CC cells.

XX Sequence 247 AA;  
 SQ Query Match 89.7%; Score 1291.5; DB 20; Length 247;  
 Best Local Similarity 90.3%; Pred. No. 1.9e-89;  
 Matches 223; Conservative 8; Mismatches 15; Indels 1; Gaps 1;

QY 5 PLHLATSFCLLSMCAQLCRPTCTCPMTPOCPGCVPLVLDGCGCCCKVCARRLGEES 64  
 DB 1 PLHLATSFCLLSMYSQLCRPAACPMTPPOCPGCVPLVLDGCGCCCKVCARRLGEES 60  
 QY 65 DHLHVCDSGGLVCPGAGPGGHCAGVCLLDEDDSCCEVNGRRYLDGETFKNCRVLCRC 124  
 DB 61 DHLHVCDSGGLVCPGAGPGGHCAGVCLLDEDDSCCEVNGRRYLDGETFKNCRVLCRC 120  
 QY 125 DDGFTCLPLCSEVDRLPSMDCPRRRIOVPGRCCPEWVCDQAVMPAIOPSAOGHQLS 183  
 DB 121 DDGFTCLPLCSEVDRLPSMDCPRRRIOVPGRCCPEWVCDQAVMPAIOPSAOGHQLS 180  
 QY 184 LVTASADAPCPMNSTAMGPGSTTCGLGIATRVSNONRFOCLEIQRRLCLSRCLASRS 243  
 DB 181 LVTASADAPCPMNSTAMGPGSTTCGLGIATRVSNONRFOCLEIQRRLCLSRCLASRS 240  
 QY 244 SSWNSAF 250  
 DB 241 GSWNSAF 247

RESULT 8  
 AAY17684  
 ID AAY17684 standard; Protein: 246 AA.  
 AC AAY17684;  
 XX 06-AUG-1999 (first entry)  
 DT Mouse WISP-2 protein SEQ ID NO:82.  
 XX  
 KW Wnt-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;  
 KW connective tissue growth factor; cancer; melanoma; arteriosclerosis;  
 KW leukaemia; lymphoid malignancy; haematopoiesis-related disorder;  
 KW tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;  
 KW kidney disorder; bone-related disorder; osteoporosis; trauma; burn;  
 KW connective tissue disorder; catabolic state; inflammation;  
 KW testicular-related disorder; angiogenesis; immunological disorder.  
 XX  
 OS Mus sp.  
 XX  
 PN WO9921998-A1.  
 XX  
 PD 06-MAY-1999.  
 DT 29-OCT-1998; 98WO-US22991.  
 XX  
 PF 14-APR-1998; 98US-0081695.  
 PR 29-OCT-1997; 97US-0063704.  
 PR 03-FEB-1998; 98US-0073612.

XX (GENENTECH INC.  
 PA Botstein DA, Cohen RL, Goddard A, Gurney AL, Hillan K;  
 PI Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WI;  
 XX WPI: 1999-337420/28.  
 DR  
 XX  
 PT New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3  
 XX  
 PS Disclosure: Page 249; 284pp; English.  
 XX  
 CC The present invention describes Wnt-1 induced secreted polypeptides,  
 CC WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2  
 CC and WISP-3 have homology to connective tissue growth factor (CTGF).  
 CC Products from the present invention can be used to treat WISP-related  
 CC disorders such as breast, ovarian, and colon cancer or melanoma. The  
 CC products can be used to treat arteriosclerosis. The products can also be  
 CC used to treat other diseases e.g. benign and malignant tumours,  
 CC leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,  
 CC hypothalamic and other glandular, macrophagal, epithelial, stromal, and  
 CC blastocoeleic disorders, haematopoiesis-related disorders, tissue-growth  
 CC disorders, skin disorders, desmoplasia, fibrotic lesions, kidney  
 CC disorders, bone-related disorders such as osteoporosis, trauma such as  
 CC burns, incisions, and other wounds, connective tissue disorders,  
 CC catabolic states, testicular-related disorders, and inflammatory,  
 CC angiogenic and immunologic disorders including arteriosclerosis. The  
 CC products can also be used for detection and diagnosis especially of  
 CC individuals with neoplastic cell growth or proliferation. The products  
 CC can be used in the production of transgenic or knock-out animals.  
 CC Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing  
 CC cells.  
 XX  
 SQ Sequence 246 AA:  
 Query Match 89.2%; Score 1284.5; DB 20; Length 246;  
 Best Local Similarity 90.2%; Pred. No. 6.5e-89;  
 Matches 222; Conservative 8; Mismatches 15; Indels 1; Gaps 1;  
 QY 6 LHLATSFCLLSMWCAQLCRTPCTCPMTPOCGVPLVLDGCGCCVCAARRLGESCD 65  
 DB 1 LHLALISFLCLLSMWYSOLCPAPACPMTPPOCGVPLVLDGCGCCVCAARRLGESCD 60  
 QY 66 LHVCDPSGGLVQCPAGPGHGAVALDDEDDGSCVNRRLYDGTFRPNCVLCRCDD 125  
 DB 61 LHVCDPSGGLVQCPAGPGHGAVALDDEDDGSCVNRRLYDGTFRPNCVLCRCDD 120  
 QY 126 GFTCLPLCSEDEVRLPSMCCPRPRRIQVPGRCCEWVCDQAVMOPAIQPSAOGHQLSL 184  
 DB 121 GFTCLPLCSEDEVRLPSMCCPRPRRIQVPGRCCEWVCDQAVMOPAIQPSAOGHQLSL 180  
 QY 185 VTPASADAPCPNMSTAMGPGSTTCGIGIATRVSNQNRFCQLEIQRRLCPRLCLARSHS 244  
 DB 181 VTPASADGCPNMSTAMGPGSTTCGIGIATRVSNQNRFCQLEIQRRLCLSRCLASRSHS 240  
 QY 245 SWNSAF 250  
 DB 241 SWNSAF 246  
 RESULT 9  
 AA117685  
 ID AA117685 standard; Protein; 245 AA.  
 XX  
 AC AA117685;  
 XX  
 DT 06-AUG-1999 (first entry)  
 XX  
 DE Mouse WISP-2 protein SEQ ID NO:83.  
 XX  
 XX WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;  
 KM connective tissue growth factor; cancer; melanoma; arteriosclerosis;  
 KM leukaemia; lymphoid malignancy; haematopoiesis-related disorder;

KM tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;  
 KM kidney disorder; bone-related disorder; osteoporosis; trauma; burn;  
 KM connective tissue disorder; catabolic state; inflammation;  
 KM testicular-related disorder; angiogenesis; immunological disorder.  
 XX  
 OS Mus sp.  
 XX  
 PN WO9921998-A1.  
 XX  
 XX 06-MAY-1999.  
 PD  
 XX 29-OCT-1998; 98WO-US22991.  
 PF  
 XX 14-APR-1998; 98US-0081695.  
 PR 29-OCT-1997; 97US-0063704.  
 PR 03-FEB-1998; 98US-0073612.  
 XX  
 PA (GENENTECH INC.  
 XX Botstein DA, Cohen RL, Goddard A, Gurney AL, Hillan K;  
 PI Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WI;  
 XX WPI: 1999-337420/28.  
 DR  
 XX  
 PT New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3  
 XX  
 PS Disclosure: Page 249-250; 284pp; English.  
 XX  
 CC The present invention describes Wnt-1 induced secreted polypeptides,  
 CC WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2  
 CC and WISP-3 have homology to connective tissue growth factor (CTGF).  
 CC Products from the present invention can be used to treat WISP-related  
 CC disorders such as breast, ovarian, and colon cancer or melanoma. The  
 CC products can be used to treat arteriosclerosis. The products can also be  
 CC used to treat other diseases e.g. benign and malignant tumours,  
 CC leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,  
 CC hypothalamic and other glandular, macrophagal, epithelial, stromal, and  
 CC blastocoeleic disorders, haematopoiesis-related disorders, tissue-growth  
 CC disorders, skin disorders, desmoplasia, fibrotic lesions, kidney  
 CC disorders, bone-related disorders such as osteoporosis, trauma such as  
 CC burns, incisions, and other wounds, connective tissue disorders,  
 CC catabolic states, testicular-related disorders, and inflammatory,  
 CC angiogenic and immunologic disorders including arteriosclerosis. The  
 CC products can also be used for detection and diagnosis especially of  
 CC individuals with neoplastic cell growth or proliferation. The products  
 CC can be used in the production of transgenic or knock-out animals.  
 CC Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing  
 CC cells.  
 XX  
 SQ Sequence 245 AA:  
 Query Match 88.9%; Score 1280.5; DB 20; Length 245;  
 Best Local Similarity 90.2%; Pred. No. 1.3e-88;  
 Matches 221; Conservative 8; Mismatches 15; Indels 1; Gaps 1;  
 QY 7 LHLATSFCLLSMWCAQLCRTPCTCPMTPOCGVPLVLDGCGCCVCAARRLGESCDH 66  
 DB 1 LHLALISFLCLLSMWYSOLCPAPACPMTPPOCGVPLVLDGCGCCVCAARRLGESCDH 60  
 QY 67 LHVCDPSGGLVQCPAGPGHGAVALDDEDDGSCVNRRLYDGTFRPNCVLCRCDDG 126  
 DB 61 LHVCDPSGGLVQCPAGPGHGAVALDDEDDGSCVNRRLYDGTFRPNCVLCRCDDG 120  
 QY 127 GFTCLPLCSEDEVRLPSMCCPRPRRIQVPGRCCEWVCDQAVMOPAIQPSAOGHQLSLV 185  
 DB 121 GFTCLPLCSEDEVRLPSMCCPRPRRIQVPGRCCEWVCDQAVMOPAIQPSAOGHQLSLV 180  
 QY 186 TTPASADAPCPNMSTAMGPGSTTCGIGIATRVSNQNRFCQLEIQRRLCPRLCLARSHS 245  
 DB 181 TTPASADGCPNMSTAMGPGSTTCGIGIATRVSNQNRFCQLEIQRRLCLSRCLASRSHS 240  
 QY 246 WNSAF 250  
 XXXX

Db 241 WNSAF 245

RESULT 10

AA17686

ID AA17686 standard; Protein: 244 AA.

XX

AC AA17686;

XX

DT 06-AUG-1999 (first entry)

XX

DE Mouse WISP-2 protein SEQ ID NO:84.

XX

KM WNT-1 induced secreted protein: WISP-1; WISP-2; WISP-3; CTGF; tumour;

KM connective tissue growth factor; cancer; melanoma; arteriosclerosis;

KM leukaemia; lymphoid malignancy; haematopoiesis-related disorder;

KM tissue growth disorder; skin disorder; desmoplasia; fibrotic lesion;

KM kidney disorder; bone-related disorder; osteoporosis; trauma; burn;

KM connective tissue disorder; catabolic state; inflammation;

KM testicular-related disorder; angiogenesis; immunological disorder.

XX

OS Mus sp.

XX

PN W09921998-A1.

XX

PD 06-MAY-1999.

XX

PE 29-OCT-1998; 98WO-US22991.

XX

PR 14-APR-1998; 98US-0081695.

PR 29-OCT-1997; 97US-0063704.

PR 03-FEB-1998; 98US-0073612.

XX

PA (GETH ) GENENTECH INC.

XX

PI Botstein DA, Cohen RL, Goddard A, Gurney AL, Hillan K;

PI Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WI;

XX

DR WPI; 1999-337420/28.

XX

FT New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3

XX

PS Disclosure; Page 250-251; 284pp; English.

XX

CC The present invention describes Wnt-1 induced secreted polypeptides,

CC WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2

CC and WISP-3 have homology to connective tissue growth factor (CTGF).

CC Products from the present invention can be used to treat WISP-related

CC disorders such as breast, ovarian, and colon cancer or melanoma. The

CC products can be used to treat arteriosclerosis. The products can also be

CC used to treat other diseases e.g. benign and malignant tumours,

CC leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,

CC hypothalamic and other glandular, macrophagal, epithelial, stromal, and

CC blastocoeleic disorders, haematopoiesis-related disorders, tissue-growth

CC disorders, skin disorders, desmoplasia, fibrotic lesions, kidney

CC disorders, bone-related disorders such as osteoporosis, trauma such as

CC burns, incisions, and other wounds, connective tissue disorders,

CC catabolic states, testicular-related disorders, and inflammatory,

CC angiogenic and immunologic disorders including arteriosclerosis. The

CC products can also be used for detection and diagnosis especially of

CC individuals with neoplastic cell growth or proliferation. The products

CC can be used in the production of transgenic or knock-out animals.

CC Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing

CC cells.

XX

XX

SQ Sequence 244 AA;

Query Match 88.6%; Score 1276.5; DB 20; Length 244;

Best Local Similarity 90.2%; Pred. No. 2,6e-88;

Matches 220; Conservative 8; Mismatches 15; Indels 1; Gaps 1;

QY 8 HLATSFCLISWVCAQLCRPTGTCPCWTPPOCCPGVPLVLDGCCCCRYCARRLGESCDEL 67

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1 HLATISFLCLISWVYSQLCPAPACAPWTPPOCCPGVPLVLDGCCCCRYCARRLGESCDEL 60

QY 68 HVCDPSGGLVCCQAPAGGAGHCAVCLLDEDDSCSCVNRRTYLDGTFEPKNCRVLCRCDDGG 127

Db 61 HVCDPSGGLVCCQAPAGGAGHCAVCLLDEDDSCSCVNRRTYLDGTFEPKNCRVLCRCDDGG 120

QY 128 FTCLPLCSEDEVRLPSPMCPPRRRIQVPGKCCPEWVCCDGV-TPAIQSTAGHQLSLALT 186

Db 121 FTCLPLCSEDEVRLPSPMCPPRRRIQVPGKCCPEWVCCDGV-TPAIQSTAGHQLSLALT 180

QY 187 PASADAPCPMWSWAMGFCSTTCGLGIATRVSNQNRFCOLEIQRRLCPRLCLARSHSSW 246

Db 181 PASADGFCPPMWSWAMGFCSTTCGLGIATRVSNQNRFCOLEIQRRLCLARSHSSW 240

QY 247 NSAF 250

Db 241 NSAF 244

RESULT 11

AA17687

ID AA17687 standard; Protein: 243 AA.

XX

AC AA17687;

XX

DT 06-AUG-1999 (first entry)

XX

DE Mouse WISP-2 protein SEQ ID NO:85.

XX

KM WNT-1 induced secreted protein: WISP-1; WISP-2; WISP-3; CTGF; tumour;

KM connective tissue growth factor; cancer; melanoma; arteriosclerosis;

KM leukaemia; lymphoid malignancy; haematopoiesis-related disorder;

KM tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;

KM kidney disorder; bone-related disorder; osteoporosis; trauma; burn;

KM connective tissue disorder; catabolic state; inflammation;

KM testicular-related disorder; angiogenesis; immunological disorder.

XX

OS Mus sp.

XX

PN W09921998-A1.

XX

PD 06-MAY-1999.

XX

PE 29-OCT-1998; 98WO-US22991.

XX

PR 14-APR-1998; 98US-0081695.

PR 29-OCT-1997; 97US-0063704.

PR 03-FEB-1998; 98US-0073612.

XX

PA (GETH ) GENENTECH INC.

XX

PI Botstein DA, Cohen RL, Goddard A, Gurney AL, Hillan K;

PI Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WI;

XX

DR WPI; 1999-337420/28.

XX

FT New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3

XX

PS Disclosure; Page 251-252; 284pp; English.

XX

CC The present invention describes Wnt-1 induced secreted polypeptides,

CC WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2

CC and WISP-3 have homology to connective tissue growth factor (CTGF).

CC Products from the present invention can be used to treat WISP-related

CC disorders such as breast, ovarian, and colon cancer or melanoma. The

CC products can be used to treat arteriosclerosis. The products can also be

CC used to treat other diseases e.g. benign and malignant tumours,

CC leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,

CC hypothalamic and other glandular, macrophagal, epithelial, stromal, and

CC blastocoeleic disorders, haematopoiesis-related disorders, tissue-growth

CC disorders, skin disorders, desmoplasia, fibrotic lesions, kidney

CC disorders, bone-related disorders such as osteoporosis, trauma such as

CC burns, incisions, and other wounds, connective tissue disorders,

CC catabolic states, testicular-related disorders, and inflammatory,  
 CC angiogenic and immunologic disorders including arteriosclerosis. The  
 CC products can also be used for detection and diagnosis especially of  
 CC individuals with neoplastic cell growth or proliferation. The products  
 CC can be used in the production of transgenic or knock-out animals.  
 CC Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing  
 CC cells.

XX Sequence 243 AA:

Query Match 88.1%; Score 1268.5; DB 20; Length 243;  
 Best Local Similarity 90.1%; Pred. No. 1e-87;

Matches 219; Conservative 8; Mismatches 15; Indels 1; Gaps 1;

OY 9 LLAISFLCLLSMWCAQLCTPCTCPMTPOCPGVPVLDGCGCCVCARRLGESCDHLH 68  
 DB 1 LLAISFLCLLSMWVSQLCPAPCACPMTPPOCPGVPVLDGCGCCVCARRLGESCDHLH 60  
 OY 69 VCDPSGGLVCGQAGGAGGAVCLLDEDDGCEVNGRRYLDGTFKPCRVLCRCDDGGF 128  
 DB 61 VCDPSGGLVCGQAGGAGGAVCLFEEDGSCVENGRYLDGTFKPCRVLCRCDDGGF 120  
 OY 129 TCLPLCSEEDVRLPSMDCPRPRRIQVPGRCCEPVVCDQAVWQPAIQSSAGHLSALVTP 187  
 DB 121 TCLPLCSEEDVRLPSMDCPRPRRIQVPGRCCEPVVCDQAVWQPAIQSSAGHLSALVTP 180  
 OY 188 ASADAPCPNMTSTAMGFCSTTCGIGIATRVSNONRFCEIQRRLCLPRCLAARSHSSWN 247  
 DB 181 ASADGCPNMTSTAMGFCSTTCGIGIATRVSNONRFCEIQRRLCLPRCLAARSHSSWN 240

OY 248 SAF 250  
 DB 241 SAF 243

# RESULT 12

ID AAY17688 standard; Protein: 242 AA.

AC AAY17688;

DT 06-AUG-1999 (first entry)

DE Mouse WISP-2 protein SEQ ID NO:86.

XX WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;  
 KW connective tissue growth factor; cancer; melanoma; arteriosclerosis;  
 KW leukaemia; lymphoid malignancy; hematopoiesis-related disorder;  
 KW tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;  
 KW kidney disorder; bone-related disorder; osteoporosis; trauma; burn;  
 KW connective tissue disorder; catabolic state; inflammation;  
 KW testicular-related disorder; angiogenesis; immunological disorder.

OS Mus sp.

PN W09921998-A1.

PD 06-MAY-1999.

PF 29-OCT-1998; 98MO-US22991.

PR 14-APR-1998; 98US-0081695.

PR 29-OCT-1997; 97US-0063704.

PR 03-FEB-1998; 98US-0073612.

PA (GENH ) GENENTECH INC.

PI Botstein DA, Cohen RL, Goddard A, Gurney AL, Hillan K;

XX Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WI;

XX WPI; 1999-337420/28.

PT New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3

XX Disclousure; Page 252-253; 284pp; English.

PS The present invention describes Wnt-1 induced secreted polypeptides,  
 XX WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2  
 CC and WISP-3 have homology to connective tissue growth factor (CTGF).  
 CC Products from the present invention can be used to treat WISP-related  
 CC disorders such as breast, ovarian, and colon cancer or melanoma. The  
 CC products can be used to treat arteriosclerosis. The products can also be  
 CC used to treat other diseases e.g. benign and malignant tumours,  
 CC leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,  
 CC hypothalamic and other glanular, macropagal, epithelial, stromal, and  
 CC blastococellic disorders, haematopolesis-related disorders, tissue-growth  
 CC disorders, skin disorders, desmoplasia, fibrotic lesions, kidney  
 CC disorders, bone-related disorders such as osteoporosis, trauma such as  
 CC burns, incisions, and other wounds, connective tissue disorders,  
 CC catabolic states, testicular-related disorders, and inflammatory,  
 CC angiogenic and immunologic disorders including arteriosclerosis. The  
 CC products can also be used for detection and diagnosis especially of  
 CC individuals with neoplastic cell growth or proliferation. The products  
 CC can be used in the production of transgenic or knock-out animals.  
 CC Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing  
 CC cells.

XX Sequence 242 AA:

Query Match 87.8%; Score 1264.5; DB 20; Length 242;  
 Best Local Similarity 90.1%; Pred. No. 2e-87;

Matches 218; Conservative 8; Mismatches 15; Indels 1; Gaps 1;

OY 10 LATSFLCLLSMWCAQLCTPCTCPMTPOCPGVPVLDGCGCCVCARRLGESCDHLH 69  
 DB 1 LATSFLCLLSMWVSQLCPAPCACPMTPPOCPGVPVLDGCGCCVCARRLGESCDHLH 60

OY 70 CDPSSGLVCGQAGGAGGAVCLLDEDDGCEVNGRRYLDGTFKPCRVLCRCDDGGF 129  
 DB 61 CDPSSGLVCGQAGGAGGAVCLFEEDGSCVENGRYLDGTFKPCRVLCRCDDGGF 120

OY 130 CLPLCSEEDVRLPSMDCPRPRRIQVPGRCCEPVVCDQAVWQPAIQSSAGHLSALVTPA 188  
 DB 121 CLPLCSEEDVRLPSMDCPRPRRIQVPGRCCEPVVCDQAVWQPAIQSSAGHLSALVTPA 180

OY 189 SADAPCPNMTSTAMGFCSTTCGIGIATRVSNONRFCEIQRRLCLPRCLAARSHSSWN 248  
 DB 181 SADGCPNMTSTAMGFCSTTCGIGIATRVSNONRFCEIQRRLCLPRCLAARSHSSWN 240

OY 249 AF 250  
 DB 241 AF 242

# RESULT 13

ID AAY17689 standard; Protein: 241 AA.

AC AAY17689;

DT 06-AUG-1999 (first entry)

DE Mouse WISP-2 protein SEQ ID NO:87.

XX WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;  
 KW connective tissue growth factor; cancer; melanoma; arteriosclerosis;  
 KW leukaemia; lymphoid malignancy; hematopoiesis-related disorder;  
 KW tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;  
 KW kidney disorder; bone-related disorder; osteoporosis; trauma; burn;  
 KW connective tissue disorder; catabolic state; inflammation;  
 KW testicular-related disorder; angiogenesis; immunological disorder.

OS Mus sp.

PN W09921998-A1.

PD 06-MAY-1999.  
 XX 29-OCT-1998; 98WO-US22991.  
 XX 14-APR-1998; 98US-0081695.  
 PR 29-OCT-1997; 97US-0063704.  
 PR 03-FEB-1998; 98US-0073612.  
 XX (GETH ) GENENTECH INC.  
 XX Botstein DA, Cohen RL, Goddard A, Gurney AL, Hillan K;  
 PI Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WI;  
 DR WPI: 1999-337420/28.  
 XX New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3  
 PT  
 XX Disclosure: Page 253-254; 284pp; English.  
 PS  
 XX The present invention describes Wnt-1 induced secreted polypeptides,  
 CC WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2  
 CC and WISP-3 have homology to connective tissue growth factor (CTGF).  
 CC Products from the present invention can be used to treat WISP-related  
 CC disorders such as breast, ovarian, and colon cancer or melanoma. The  
 CC products can be used to treat arteriosclerosis. The products can also be  
 CC used to treat other diseases e.g. benign and malignant tumours,  
 CC leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,  
 CC hypothalamic and other glandular, macrophagal, epithelial, stromal, and  
 CC blastocoele disorders, haematopoiesis-related disorders, tissue-growth  
 CC disorders, skin disorders, desmoplasia, fibrotic lesions, kidney  
 CC burns, incisions, bone-related disorders such as osteoporosis, trauma such as  
 CC catabolic states, testicular-related disorders, and inflammatory,  
 CC angiogenic and immunologic disorders including arteriosclerosis. The  
 CC products can also be used for detection and diagnosis especially of  
 CC individuals with neoplastic cell growth or proliferation. The products  
 CC can be used in the production of transgenic or knock-out animals.  
 CC Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing  
 CC cells.  
 XX  
 XX Sequence 241 AA:  
 SQ  
 Query Match 87.5%; Score 1260.5; DB 20; Length 241;  
 Best Local Similarity 90.0%; Pred. No. 4e-87;  
 Matches 217; Conservative 8; Mismatches 15; Indels 1; Gaps 1;

DT 06-AUG-1999 (first entry)  
 XX  
 DE Mouse WISP-2 protein SEQ ID NO:89.  
 XX  
 XX WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;  
 KW connective tissue growth factor; cancer; melanoma; arteriosclerosis;  
 KW leukaemia; lymphoid malignancy; haematopoiesis-related disorder;  
 KW tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;  
 KW kidney disorder; bone-related disorder; osteoporosis; trauma; burn;  
 KW connective tissue disorder; catabolic state; inflammation;  
 KW testicular-related disorder; angiogenesis; immunological disorder.  
 XX  
 OS Mus sp.  
 XX  
 XX W09921998-A1.  
 XX  
 PD 06-MAY-1999.  
 XX  
 XX 29-OCT-1998; 98WO-US22991.  
 XX  
 XX 14-APR-1998; 98US-0081695.  
 PR 29-OCT-1997; 97US-0063704.  
 PR 03-FEB-1998; 98US-0073612.  
 XX (GETH ) GENENTECH INC.  
 XX Botstein DA, Cohen RL, Goddard A, Gurney AL, Hillan K;  
 PI Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WI;  
 DR WPI: 1999-337420/28.  
 XX New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3  
 PT  
 XX Disclosure: Page 255-256; 284pp; English.  
 PS  
 XX The present invention describes Wnt-1 induced secreted polypeptides,  
 CC WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2  
 CC and WISP-3 have homology to connective tissue growth factor (CTGF).  
 CC Products from the present invention can be used to treat WISP-related  
 CC disorders such as breast, ovarian, and colon cancer or melanoma. The  
 CC products can be used to treat arteriosclerosis. The products can also be  
 CC used to treat other diseases e.g. benign and malignant tumours,  
 CC leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,  
 CC hypothalamic and other glandular, macrophagal, epithelial, stromal, and  
 CC blastocoele disorders, haematopoiesis-related disorders, tissue-growth  
 CC disorders, skin disorders, desmoplasia, fibrotic lesions, kidney  
 CC disorders, bone-related disorders such as osteoporosis, trauma such as  
 CC burns, incisions, and other wounds, connective tissue disorders,  
 CC catabolic states, testicular-related disorders, and inflammatory,  
 CC angiogenic and immunologic disorders including arteriosclerosis. The  
 CC products can also be used for detection and diagnosis especially of  
 CC individuals with neoplastic cell growth or proliferation. The products  
 CC can be used in the production of transgenic or knock-out animals.  
 CC Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing  
 CC cells.  
 XX  
 XX Sequence 239 AA:  
 SQ  
 Query Match 87.3%; Score 1257.5; DB 20; Length 239;  
 Best Local Similarity 90.4%; Pred. No. 6.7e-87;  
 Matches 216; Conservative 8; Mismatches 14; Indels 1; Gaps 1;

OY 192 APCPNMSTAMGPGSTTCGIGIATRVSNQNRFCOLEIQRRLCLPRPCLARSHSSWNSAF 250  
DB 181 GPCPNMSTAMGPGSTTCGIGIATRVSNQNRFCOLEIQRRLCLSRPCLARSHSSWNSAF 239

RESULT 15  
AAV17690  
ID AAV17690 standard; Protein; 240 AA.  
XX  
AC AAV17690;  
XX  
DT 06-AUG-1999 (first entry)  
XX  
DE Mouse WISP-2 protein SEQ ID NO:88.  
XX  
KM WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;  
KM connective tissue growth factor; cancer; melanoma; arteriosclerosis;  
KM leukaemia; lymphoid malignancy; haematopoiesis-related disorder;  
KM tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;  
KM kidney disorder; bone-related disorder; osteoporosis; trauma; burn;  
KM connective tissue disorder; catabolic state; inflammation;  
KM testicular-related disorder; angiogenesis; immunological disorder.  
XX  
OS Mus sp.  
XX  
PN WO9921998-A1.  
XX  
PD 06-MAY-1999.  
XX  
PF 29-OCT-1998; 98WO-US22991.  
XX  
PR 14-APR-1998; 98US-0081695.  
PR 29-OCT-1997; 97US-0063704.  
PR 03-FEB-1998; 98US-0073612.  
XX  
PA (GETH ) GENENTECH INC.  
XX  
PI Botstein DA, Cohen RL, Goddard A, Gurney AL, Hillan K;  
PI Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WI;  
XX  
DR WPI: 1999-337420/28.  
XX  
PT New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3  
XX  
PS Disclosure: Page 254-255; 284pp; English.  
XX  
CC The present invention describes Wnt-1 induced secreted polypeptides,  
CC WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2  
CC and WISP-3 have homology to connective tissue growth factor (CTGF).  
CC Products from the present invention can be used to treat WISP-related  
CC disorders such as breast, ovarian, and colon cancer or melanoma. The  
CC products can be used to treat arteriosclerosis. The products can also be  
CC used to treat other diseases e.g. benign and malignant tumours,  
CC leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,  
CC hypothalamic and other glandular, macrophagal, epithelial, stromal, and  
CC blastocoealic disorders, haematopoiesis-related disorders, tissue-growth  
CC disorders, skin disorders, desmoplasia, fibrotic lesions, kidney  
CC disorders, bone-related disorders such as osteoporosis, trauma such as  
CC burns, incisions, and other wounds, connective tissue disorders,  
CC catabolic states, testicular-related disorders, and inflammatory,  
CC angiogenic and immunologic disorders including arteriosclerosis. The  
CC products can also be used for detection and diagnosis especially of  
CC individuals with neoplastic cell growth or proliferation. The products  
CC can be used in the production of transgenic or knock-out animals.  
CC Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing  
CC cells.  
XX  
SQ Sequence 240 AA;

Query Match 87.3%; Score 1257.5; DB 20; Length 240;  
Best Local Similarity 90.4%; Pred. No. 6.7e-87;  
Matches 216; Conservative 8; Mismatches 14; Indels 1; Gaps 1;

OY 13 SFCLISWVCAQLCRTPTCTCPWTPPQCPOGVPLVLDGCGCGKVCARLGECDHLHVCDP 72  
DB 2 SFLGISWVYSQLCAPACACPMTPPQCPRPGVPLVLDGCGCGKVCARLGECDHLHVCDP 61  
OY 73 SGLVCGPGAGPGGHCANCLDDEBDSCEVNGRRYIDGTFKPNCRVLCRDDGGFTCLP 132  
DB 62 SGLVCGPGAGPGGHCANCLDDEBDSCEVNGRRYIDGTFKPNCRVLCRDDGGFTCLP 121  
OY 133 LCEVDVRLPSMDCPRPRRIQVPGRCCPEWVCDQAVWQPAIQPSAOGHOLSALVTPASAD 191  
DB 122 LCEVDVRLPSMDCPRPRRIQVPGRCCPEWVCDQAVWQPAIQPSAOGHOLSALVTPASAD 181  
OY 192 APCPNMSTAMGPGSTTCGIGIATRVSNQNRFCOLEIQRRLCLPRPCLARSHSSWNSAF 250  
DB 182 GPCPNMSTAMGPGSTTCGIGIATRVSNQNRFCOLEIQRRLCLSRPCLARSHSSWNSAF 240

Search completed: July 25, 2003, 12:34:59  
Job time : 85 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 25, 2003, 12:33:35 ; Search time 29 Seconds  
(without alignments)  
364.749 Million cell updates/sec

Title: US-10-010-408-2  
Perfect score: 1440  
Sequence: 1 MRGSPHLHLATSLFLISM.....LCPLRPLAARSHSNNSAF 250

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: Issued\_Patents\_AA.\*  
2: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep.\*  
3: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep.\*  
4: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep.\*  
5: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep.\*  
6: /cgn2\_6/ptodata/1/1aa/6CTUS.COMB.pep.\*  
7: /cgn2\_6/ptodata/1/1aa/6CTUS.COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score  | Query Match | Length | ID               | Description       |
|------------|--------|-------------|--------|------------------|-------------------|
| 1          | 1308.5 | 90.9        | 251    | US-09-182-145-20 | Sequence 20, Appl |
| 2          | 1303.5 | 90.5        | 250    | US-09-182-145-78 | Sequence 78, Appl |
| 3          | 1298.5 | 90.2        | 249    | US-09-182-145-79 | Sequence 79, Appl |
| 4          | 1292.5 | 89.8        | 248    | US-09-182-145-80 | Sequence 80, Appl |
| 5          | 1291.5 | 89.7        | 247    | US-09-182-145-81 | Sequence 81, Appl |
| 6          | 1284.5 | 89.2        | 246    | US-09-182-145-82 | Sequence 82, Appl |
| 7          | 1280.5 | 88.9        | 245    | US-09-182-145-83 | Sequence 83, Appl |
| 8          | 1276.5 | 88.6        | 244    | US-09-182-145-84 | Sequence 84, Appl |
| 9          | 1268.5 | 88.1        | 243    | US-09-182-145-85 | Sequence 85, Appl |
| 10         | 1264.5 | 87.8        | 242    | US-09-182-145-86 | Sequence 86, Appl |
| 11         | 1260.5 | 87.5        | 241    | US-09-182-145-87 | Sequence 87, Appl |
| 12         | 1257.5 | 87.3        | 239    | US-09-182-145-88 | Sequence 88, Appl |
| 13         | 1257.5 | 87.3        | 240    | US-09-182-145-89 | Sequence 89, Appl |
| 14         | 1253.5 | 87.0        | 238    | US-09-182-145-90 | Sequence 90, Appl |
| 15         | 1247.5 | 86.6        | 237    | US-09-182-145-91 | Sequence 91, Appl |
| 16         | 1243.5 | 86.4        | 236    | US-09-182-145-92 | Sequence 92, Appl |
| 17         | 1234.5 | 85.7        | 235    | US-09-182-145-93 | Sequence 93, Appl |
| 18         | 1232.5 | 85.6        | 234    | US-09-182-145-94 | Sequence 94, Appl |
| 19         | 1228.5 | 85.3        | 233    | US-09-182-145-95 | Sequence 95, Appl |
| 20         | 1224.5 | 85.0        | 232    | US-09-182-145-96 | Sequence 96, Appl |
| 21         | 1219.5 | 84.7        | 231    | US-09-182-145-97 | Sequence 97, Appl |
| 22         | 1217.5 | 84.5        | 229    | US-09-182-145-99 | Sequence 99, Appl |
| 23         | 1217.5 | 84.5        | 230    | US-09-182-145-98 | Sequence 98, Appl |
| 24         | 1216.5 | 84.5        | 228    | US-09-182-145-19 | Sequence 19, Appl |
| 25         | 1064   | 73.9        | 250    | US-09-182-145-16 | Sequence 16, Appl |
| 26         | 1059   | 73.5        | 249    | US-09-182-145-57 | Sequence 57, Appl |
| 27         | 1054   | 73.2        | 248    | US-09-182-145-57 | Sequence 57, Appl |

|    |      |      |     |                  |                   |
|----|------|------|-----|------------------|-------------------|
| 28 | 1048 | 72.8 | 247 | US-09-182-145-58 | Sequence 58, Appl |
| 29 | 1047 | 72.7 | 246 | US-09-182-145-59 | Sequence 59, Appl |
| 30 | 1043 | 72.4 | 243 | US-09-182-145-62 | Sequence 62, Appl |
| 31 | 1043 | 72.4 | 244 | US-09-182-145-61 | Sequence 61, Appl |
| 32 | 1043 | 72.4 | 245 | US-09-182-145-60 | Sequence 60, Appl |
| 33 | 1035 | 71.9 | 242 | US-09-182-145-63 | Sequence 63, Appl |
| 34 | 1031 | 71.6 | 241 | US-09-182-145-64 | Sequence 64, Appl |
| 35 | 1027 | 71.3 | 240 | US-09-182-145-65 | Sequence 65, Appl |
| 36 | 1025 | 71.2 | 238 | US-09-182-145-67 | Sequence 67, Appl |
| 37 | 1025 | 71.2 | 239 | US-09-182-145-66 | Sequence 66, Appl |
| 38 | 1021 | 70.9 | 236 | US-09-182-145-69 | Sequence 69, Appl |
| 39 | 1021 | 70.9 | 237 | US-09-182-145-68 | Sequence 68, Appl |
| 40 | 1017 | 70.6 | 235 | US-09-182-145-70 | Sequence 70, Appl |
| 41 | 1008 | 70.0 | 234 | US-09-182-145-71 | Sequence 71, Appl |
| 42 | 1004 | 69.7 | 233 | US-09-182-145-72 | Sequence 72, Appl |
| 43 | 1000 | 69.4 | 232 | US-09-182-145-73 | Sequence 73, Appl |
| 44 | 997  | 69.2 | 230 | US-09-182-145-75 | Sequence 75, Appl |
| 45 | 997  | 69.2 | 231 | US-09-182-145-74 | Sequence 74, Appl |

ALIGNMENTS

RESULT 1  
US-09-182-145-20  
Sequence 20, Application US/09182145B  
Patent No. 6387657  
GENERAL INFORMATION:  
APPLICANT: Botstein, David A.  
APPLICANT: Cohen, Robert  
APPLICANT: Goddard, Audrey L.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: Lawrence, David A.  
APPLICANT: Levine, Arnold J.  
APPLICANT: Pennica, Diane  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: WSP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME  
FILE REFERENCE: P1176R2  
CURRENT APPLICATION NUMBER: US/09/182,145B  
EARLIER FILING DATE: 1998-10-29  
EARLIER APPLICATION NUMBER: US 60/063,704  
EARLIER FILING DATE: 1997-10-29  
EARLIER APPLICATION NUMBER: US 60/073,612  
EARLIER FILING DATE: 1998-02-04  
EARLIER APPLICATION NUMBER: US 60/081,695  
EARLIER FILING DATE: 1998-04-14  
NUMBER OF SEQ ID NOS: 156  
SEQ ID NO 20  
LENGTH: 251  
TYPE: PRT  
ORGANISM: Mus musculus  
US-09-182-145-20  
Query Match 90.9% Score 1308.5 DB 4: Length 251:  
Best Local Similarity 90.0% Pred. No. 1.1e-105; Indels 1; Gaps 1;  
Matches 226; Conservative 9; Mismatches 15;

QY 1 MRGSPHLHLATSLFLISMCAQLCRPTCTPCPPGCVPLVLDGCGCCVCARRL 60  
DB 1 MRGSPHLHLATSLFLISMVYSQLCPAPACPMTPPCPGVPLVLDGCGCCVCARRL 60  
QY 61 GESCDHLAVCDPSGGLVCGPAGGAGGAGVCLLDDGSGCVRNRRYLDGTFEPNRYL 120  
DB 61 GESCDHLAVCDPSGGLVCGPAGGAGGAGVCLLDDGSGCVRNRRYLDGTFEPNRYL 120  
QY 121 CRCDDGGTCLPLCSEEDVRLPSMPCPRKRLOVGGKCCPEVNCOGV-TPAIOGSTOGH 179  
DB 121 CRCDDGGTCLPLCSEEDVRLPSMPCPRKRLOVGGKCCPEVNCOGV-TPAIOGSTOGH 180  
QY 180 QLSALVTPASADAPCPMNSTAMGPCSTTCGIGIATRVSNONRFOLEIQRRLCLPRCLA 239

Db 181 QLSALVTPASADGCPGPNMSTAMGPGCTTCGLGATATRVSNONRFOLEIQRRLCLSRPCLA 240  
QY 240 ARSHSWSNAF 250  
Db 241 SRSHGSWSNAF 251

## RESULT 2

US-09-182-145-78  
Sequence 78, Application US/09182145B  
Patent No. 6387657

## GENERAL INFORMATION:

APPLICANT: Botstein, David A.  
APPLICANT: Cohen, Robert  
APPLICANT: Goddard, Audrey  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: Lawrence, David A.  
APPLICANT: Levine, Arnold J.  
APPLICANT: Pennica, Diane  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME  
FILE REFERENCE: P1176R2  
CURRENT APPLICATION NUMBER: US/09/182,145B  
EARLIER FILING DATE: 1998-10-29  
EARLIER APPLICATION NUMBER: US 60/063,704  
EARLIER FILING DATE: 1997-10-29  
EARLIER APPLICATION NUMBER: US 60/073,612  
EARLIER FILING DATE: 1998-02-04  
EARLIER APPLICATION NUMBER: US 60/081,695  
EARLIER FILING DATE: 1998-04-14  
NUMBER OF SEQ ID NOS: 156  
SEQ ID NO 78  
LENGTH: 250  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-182-145-78

Query Match  
Best Local Similarity 90.0%; Score 1303.5; DB 4; Length 250;  
Matches 225; Conservative 9; Mismatches 15; Indels 1; Gaps 1;

QY 2 RSPPLHLATSFLLCLLSWYCAQLCRTPCTCPMTPOCGVPLVLDGCGCCVCAARRLG 61  
Db 1 RSNPLHLATSFLLCLLSWYCAQLCRTPCTCPMTPOCGVPLVLDGCGCCVCAARRLG 60  
QY 62 ESCDHLHVCDSQGLVCPGAGPGHGAACLLDEDDGSCVNGRRYLDGETFRKPCRVLCR 121  
Db 61 ESCDHLHVCDSQGLVCPGAGPGHGAACLLDEDDGSCVNGRRYLDGETFRKPCRVLCR 120  
QY 122 RCDGFTCLPLCSEDEVRLPSMDCPRPKRIQVPGKCCPEMVCDOGV-TPAIORSTAGHQL 180  
Db 121 RCDGFTCLPLCSEDEVRLPSMDCPRPKRIQVPGKCCPEMVCDOGV-TPAIORSTAGHQL 180  
QY 181 LSALVTPASADGCPGPNMSTAMGPGCTTCGLGATATRVSNONRFOLEIQRRLCLSRPCLA 240  
Db 181 LSALVTPASADGCPGPNMSTAMGPGCTTCGLGATATRVSNONRFOLEIQRRLCLSRPCLA 240  
QY 241 RSHSWSNAF 250  
Db 241 RSHGSWSNAF 250

## RESULT 3

US-09-182-145-79

Sequence 79, Application US/09182145B  
Patent No. 6387657

## GENERAL INFORMATION:

APPLICANT: Botstein, David A.  
APPLICANT: Cohen, Robert  
APPLICANT: Goddard, Audrey  
APPLICANT: Gurney, Austin L.

APPLICANT: Hillan, Kenneth J.  
APPLICANT: Lawrence, David A.  
APPLICANT: Levine, Arnold J.  
APPLICANT: Pennica, Diane  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME  
FILE REFERENCE: P1176R2  
CURRENT APPLICATION NUMBER: US/09/182,145B  
EARLIER FILING DATE: 1998-10-29  
EARLIER APPLICATION NUMBER: US 60/063,704  
EARLIER FILING DATE: 1997-10-29  
EARLIER APPLICATION NUMBER: US 60/073,612  
EARLIER FILING DATE: 1998-02-04  
EARLIER APPLICATION NUMBER: US 60/081,695  
EARLIER FILING DATE: 1998-04-14  
NUMBER OF SEQ ID NOS: 156  
SEQ ID NO 79  
LENGTH: 249  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-182-145-79

Query Match  
Best Local Similarity 90.2%; Score 1298.5; DB 4; Length 249;  
Matches 224; Conservative 9; Mismatches 15; Indels 1; Gaps 1;

QY 3 GSPPLHLATSFLLCLLSWYCAQLCRTPCTCPMTPOCGVPLVLDGCGCCVCAARRLG 62  
Db 1 GNPPLHLATSFLLCLLSWYCAQLCRTPCTCPMTPOCGVPLVLDGCGCCVCAARRLG 60  
QY 63 ESCDHLHVCDSQGLVCPGAGPGHGAACLLDEDDGSCVNGRRYLDGETFRKPCRVLCR 122  
Db 61 ESCDHLHVCDSQGLVCPGAGPGHGAACLLDEDDGSCVNGRRYLDGETFRKPCRVLCR 120  
QY 123 CDDGFTCLPLCSEDEVRLPSMDCPRPKRIQVPGKCCPEMVCDOGV-TPAIORSTAGHQL 181  
Db 121 CDDGFTCLPLCSEDEVRLPSMDCPRPKRIQVPGKCCPEMVCDOGV-TPAIORSTAGHQL 180  
QY 182 SALVTPASADGCPGPNMSTAMGPGCTTCGLGATATRVSNONRFOLEIQRRLCLSRPCLA 241  
Db 181 SALVTPASADGCPGPNMSTAMGPGCTTCGLGATATRVSNONRFOLEIQRRLCLSRPCLA 240  
QY 242 SHSWSNAF 250  
Db 241 SHGSWSNAF 249

## RESULT 4

US-09-182-145-80

Sequence 80, Application US/09182145B  
Patent No. 6387657

## GENERAL INFORMATION:

APPLICANT: Botstein, David A.  
APPLICANT: Cohen, Robert  
APPLICANT: Goddard, Audrey  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: Lawrence, David A.  
APPLICANT: Levine, Arnold J.  
APPLICANT: Pennica, Diane  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME  
FILE REFERENCE: P1176R2  
CURRENT APPLICATION NUMBER: US/09/182,145B  
EARLIER FILING DATE: 1998-10-29  
EARLIER APPLICATION NUMBER: US 60/063,704  
EARLIER FILING DATE: 1997-10-29  
EARLIER APPLICATION NUMBER: US 60/073,612  
EARLIER FILING DATE: 1998-02-04  
EARLIER APPLICATION NUMBER: US 60/081,695  
EARLIER FILING DATE: 1998-04-14



```

; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 80
; LENGTH: 248
; TYPE: prt
; ORGANISM: Homo sapiens
US-09-182-145-80

```

|                       |              |                     |                |                  |
|-----------------------|--------------|---------------------|----------------|------------------|
| Query Match           | 89.8%;       | Score 1292.5;       | DB 4;          | Length 248;      |
| Best Local Similarity | 89.9%;       | Pred. No. 2.6e-104; |                |                  |
| Matches 223;          | Conservative | 9;                  | Mismatches 15; | Indels 1; Gaps 1 |

|    |  |     |
|----|--|-----|
| QY | SPILHLATLATEFLCLSMVCAQLCRPTCTCPMPRPGCPGVPLVDGCGCCVVCARRIGES        | 63  |
| Db | 1 NPLIHLIASFLCLISLMVYSQLCPAPCACPMPRPPGCPGVPLVDGCGCCVVCARRIGES      | 60  |
| QY | 64 CDHLHVCDDPSQGLVCOPGAGPGGHGAVCLLDDDSCEVNGRRYIDGETFKNCRVLCRC      | 123 |
| Db | 61 CDHLHVCDDPSQGLVCOPGAGPGSGRGAVCLFEEDDDSCENVNGRRYIDGETFKNCRVLCRC  | 120 |
| QY | 124 DDGSGFTCLPLCSEDVRLPSMDCPRPRRIOPKCKCEPWVCCOGV--TPAIOKSTAGHOLS   | 182 |
| Db | 121 DDGSGFTCLPLCSEDVRLPSMDCPRPRRIOPKCKCEPWVCCOAWQPAIOQSPASQGHOLS   | 180 |
| QY | 183 ALVTYPASADAPCPNMSTAMGPCSTTCGSLGIATRVSNONRFCOLEIQRILCLPRCLLAARS | 242 |
| Db | 181 ALVTYPASADGPCPNMSTAMGPCSTTCGSLGIATRVSNONRFOLEIQRILCLSRCLLAARS  | 240 |
| QY | 243 HSSMNSAF 250   |     |
| Db | 241 HSSMNSAF 248   |     |

## RESULT 5

```

US-09-182-145-81
: Sequence 81. Application US/09182145B
: Patent No. 6387657
:
: GENERAL INFORMATION:
:   APPLICANT: Botstein, David A.
:   APPLICANT: Cohen, Robert
:   APPLICANT: Goddard, Audrey
:   APPLICANT: Gurney, Austin L.
:   APPLICANT: Hillan, Kenneth J.
:   APPLICANT: Lawrence, David A.
:   APPLICANT: Levine, Arnold J.
:   APPLICANT: Pennica, Diane
:   APPLICANT: Roy, Margaret Ann
:   APPLICANT: Wood, William I.
:
: TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
:
: FILE REFERENCE: P1176R2
:
: CURRENT APPLICATION NUMBER: US/09/182,145B
:
: CURRENT FILING DATE: 1998-10-29
:
: EARLIER APPLICATION NUMBER: US 60/063,704
:
: EARLIER FILING DATE: 1997-10-29
:
: EARLIER APPLICATION NUMBER: US 60/073,612
:
: EARLIER FILING DATE: 1998-02-04
:
: EARLIER APPLICATION NUMBER: US 60/081,695
:
: EARLIER FILING DATE: 1998-04-14
:
: NUMBER OF SEQ ID NOS: 156
:
: SEQ ID NO 81
:
: LENGTH: 247
:
: TYPE: PRT
:
: ORGANISM: Homo sapiens
:
: US-09-182-145-81

```

|                       |              |                     |                |             |
|-----------------------|--------------|---------------------|----------------|-------------|
| Query Match           | 89.7%;       | Score 1291.5;       | DB 4;          | Length 247; |
| Best Local Similarity | 90.3%;       | Pred. No. 3.1e-104; |                |             |
| Matches 223;          | Conservative | 8;                  | Mismatches 15; | Indels 1;   |
|                       |              |                     |                | Gaps 1;     |

QY 5 PLHLHLATSLICLAWCAQOLCRPTQCEWMPRPOCGVPLVLVGGCCGCKYCARLAGEESC 64

Db 1 PLHLHLATSLICLAWVYSQGLCPAPACACWMTPPQCPGVPVLYLDGGCCGCKYCARLAGEESC 60

65 DHLHACDPSOGLVQCPAGPGGHGAVALCLDEDDGSCVENVGRRTYLDGETFKPNCBRYLCRCD 124

|    |   |
|----|---|
| Dd | <br>61 DHHVCDPSGGLYQPGAGSPGRGAVCLFEEDDGSCEVNGRRITLDGETFRNCRVLRCRD 120 |
| Qy | 125 DGGEFTCLPLCEDVDRLPSPWDCRPRIKRIQVPKCCEWYCDGV--TPAIOIRSTAGHQLSA 183 |
| Dd | 121 DGGEFTCLPLCEDVDRLPSPWDCRPRIKRIQVPKCCEWYCDVAWPAIOPSSAOGHQLSA 180   |
| Qy | 184 LVTPASADAPCPNMSTAMGPCSTTGGLGIATFVSNONRFCOLFIQRRLCIRPCLARSH 243    |
| Dd | 181 LVTPASADGPCCPMNSTAMGPCSTTGGLGIATFVSNONRFCOLFIQRRLCIRPCLARSH 240   |
| Qy | 244 SSNNNSAF 250<br>  |
| Dd | 241 GSNNSAF 247   |

## RESULT 6

```

US-09-182-145-82
: Sequence 82, Application US/09182145B
: Patent No. 6387657
:
: GENERAL INFORMATION:
: APPLICANT: Botstein, David A.
: APPLICANT: Cohen, Robert
: APPLICANT: Goddard, Audrey
: APPLICANT: Gurney, Austin L.
: APPLICANT: Hillan, Kenneth J.
: APPLICANT: Lawrence, David A.
: APPLICANT: Levine, Arnold J.
: APPLICANT: Pennica, Diane
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Wood, William T.
: TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
:
: FILE REFERENCE: P1176R2
: CURRENT APPLICATION NUMBER: US/09/182,145B
: CURRENT FILING DATE: 1998-10-29
: EARLIER APPLICATION NUMBER: US 60/063,704
: EARLIER FILING DATE: 1997-10-29
: EARLIER APPLICATION NUMBER: US 60/073,612
: EARLIER FILING DATE: 1998-02-04
: EARLIER APPLICATION NUMBER: US 60/081,695
: EARLIER FILING DATE: 1998-04-14
: NUMBER OF SEQ ID NOS: 156
:
: SEQ ID NO 82
: LENGTH: 246
: TYPE: PRT
:
: ORGANISM: Homo sapiens
US-09-182-145-82

```

|                       |              |                     |                |             |
|-----------------------|--------------|---------------------|----------------|-------------|
| Query Match           | 89.2%;       | Score 1284.5;       | DB 4;          | Length 246; |
| Best Local Similarity | 90.2%;       | Pred. No. 1.3e-103; |                |             |
| Matches 222;          | Conservative | 8;                  | Mismatches 15; | Indels 1;   |
|                       |              |                     |                | Gaps 1;     |

[illegible]

```
RESULT 7
US-09-182-145-83
; Sequence 83, Application US/09182145B
; Patent No. 6387657
; GENERAL INFORMATION:
; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/09/182,145B
; EARLIER FILING DATE: 1998-10-29
; EARLIER APPLICATION NUMBER: US 60/063,704
; EARLIER FILING DATE: 1997-10-29
; EARLIER APPLICATION NUMBER: US 60/073,612
; EARLIER FILING DATE: 1998-02-04
; EARLIER APPLICATION NUMBER: US 60/081,695
; EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 83
; TYPE: PRT
; LENGTH: 245
; ORGANISM: Homo sapiens
US-09-182-145-83
```

```
Query Match      88.9%; Score 1280.5; DB 4; Length 245;
Best Local Similarity 90.2%; Pred. No. 2.8e-103;
Matches 221; Conservative 8; Mismatches 15; Indels 1; Gaps 1;
```

```
QY 7 IHLATSFCLLSMVCACQLCRPTCTCPMTPPCCPGVPLVLDGCGCCCKVCARRLGESCDH 66
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 IHLAISFCILSMYVSQLCRPAACACPMTPPCPGVPLVLDGCGCCCKVCARRLGESCDH 60
QY 67 LHVCDPSQGLVCPGAGPGSGHGVCLLDEDDGSCVENGRYRRLDGETFFKPNCRVLCRCDDG 126
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 61 LHVCDPSQGLVCPGAGPGSGHGVCLLDEDDGSCVENGRYRRLDGETFFKPNCRVLCRCDDG 120
QY 127 GFTCLPLCSEDEVRLPSCDPRPKRIQVPGKCCPEWVCQGV-TPAIORSTAGHOLSALV 185
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 121 GFTCLPLCSEDEVRLPSCDPRPKRIQVPGKCCPEWVCQGV-TPAIORSTAGHOLSALV 180
QY 186 TPASADAPCPNMSTWAGPSTTCGLGIATRVSNQNRFCQLEIQRRLCLRPCLAASSHSS 245
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 181 TPASADGCPNMSTWAGPSTTCGLGIATRVSNQNRFCQLEIQRRLCLRPCLAASSHSS 240
QY 246 WNSAF 250
    |||||
DB 241 WNSAF 245
```

```
RESULT 8
US-09-182-145-84
; Sequence 84, Application US/09182145B
; Patent No. 6387657
; GENERAL INFORMATION:
; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
US-09-182-145-84
```

```
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/09/182,145B
; EARLIER FILING DATE: 1998-10-29
; EARLIER APPLICATION NUMBER: US 60/063,704
; EARLIER FILING DATE: 1997-10-29
; EARLIER APPLICATION NUMBER: US 60/073,612
; EARLIER FILING DATE: 1998-02-04
; EARLIER APPLICATION NUMBER: US 60/081,695
; EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 84
; TYPE: PRT
; LENGTH: 244
; ORGANISM: Homo sapiens
US-09-182-145-84
```

```
Query Match      88.6%; Score 1276.5; DB 4; Length 244;
Best Local Similarity 90.2%; Pred. No. 6.1e-103;
Matches 220; Conservative 8; Mismatches 15; Indels 1; Gaps 1;
```

```
QY 8 HLLATSFCLLSMVCACQLCRPTCTCPMTPPCCPGVPLVLDGCGCCCKVCARRLGESCDH 67
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 HLLAISFCILSMYVSQLCRPAACACPMTPPCPGVPLVLDGCGCCCKVCARRLGESCDH 60
QY 68 HVCDDPSQGLVCPGAGPGSGHGVCLLDEDDGSCVENGRYRRLDGETFFKPNCRVLCRCDDG 127
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 61 HVCDDPSQGLVCPGAGPGSGHGVCLLDEDDGSCVENGRYRRLDGETFFKPNCRVLCRCDDG 120
QY 128 FTCLPLCSEDEVRLPSCDPRPKRIQVPGKCCPEWVCQGV-TPAIORSTAGHOLSALV 186
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 121 FTCLPLCSEDEVRLPSCDPRPKRIQVPGKCCPEWVCQGV-TPAIORSTAGHOLSALV 180
QY 187 PASADAPCPNMSTWAGPSTTCGLGIATRVSNQNRFCQLEIQRRLCLRPCLAASSHSS 246
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 181 PASADGCPNMSTWAGPSTTCGLGIATRVSNQNRFCQLEIQRRLCLRPCLAASSHSS 240
QY 247 NSAF 250
    |||||
DB 241 NSAF 244
```

```
RESULT 9
US-09-182-145-85
; Sequence 85, Application US/09182145B
; Patent No. 6387657
; GENERAL INFORMATION:
; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/09/182,145B
; EARLIER FILING DATE: 1998-10-29
; EARLIER APPLICATION NUMBER: US 60/063,704
; EARLIER FILING DATE: 1997-10-29
; EARLIER APPLICATION NUMBER: US 60/073,612
; EARLIER FILING DATE: 1998-02-04
; EARLIER APPLICATION NUMBER: US 60/081,695
; EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 85
; TYPE: PRT
; LENGTH: 243
; ORGANISM: Homo sapiens
US-09-182-145-85
```





QY 193 PCPNWSTAMGECSTTCGIGIATRVSNQNRFCOLEIQRRLCLPRCLAASSHSSWNSAF 250  
Db 181 PCPNWSTAMGECSTTCGIGIATRVSNQNRFCOLEIQRRLCLSRPCLASRSHGSMNSAF 238

RESULT 15  
US-09-182-145-91  
; Sequence 91, Application US/09182145B  
; Patent No. 6387657

GENERAL INFORMATION:

; APPLICANT: Botstein, David A.  
; APPLICANT: Cohen, Robert  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth J.  
; APPLICANT: Lawrence, David A.  
; APPLICANT: Levine, Arnold J.  
; APPLICANT: Pennica, Diane  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Wood, William I.  
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: P1176R2  
; CURRENT APPLICATION NUMBER: US/09/182,145B  
; CURRENT FILING DATE: 1998-10-29  
; EARLIER APPLICATION NUMBER: US 60/063,704  
; EARLIER FILING DATE: 1997-10-29  
; EARLIER APPLICATION NUMBER: US 60/073,612  
; EARLIER FILING DATE: 1998-02-04  
; EARLIER APPLICATION NUMBER: US 60/081,695  
; EARLIER FILING DATE: 1998-04-14  
; NUMBER OF SEQ ID NOS: 156  
; SEQ ID NO 91  
; LENGTH: 237  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-182-145-91

Query Match 86.6%; Score 1247.5; DB 4; Length 237;  
Best Local Similarity 90.3%; Pred. No. 1.9e-100;  
Matches 214; Conservative 8; Mismatches 14; Indels 1; Gaps 1;

QY 15 LCLLSMYCAQLCFRTCTCPMTTPPOCPGQVPLVLDGCGCCYCARRLGSCDHLHYCDPSQ 74  
Db 1 LCLLSMYSQLCPACPCAPMTTPPOCPGQVPLVLDGCGCCYCARRLGSCDHLHYCDPSQ 60  
QY 75 GLVQPGAGPGGHAVALDDEDDGSCVEYNGRRYLDGETFKPNCRYLCRCDDGFTCLPLC 134  
Db 61 GLVQPGAGPGGHAVALDDEDDGSCVEYNGRRYLDGETFKPNCRYLCRCDDGFTCLPLC 120  
QY 135 SEDVRLPSMDCPRRRTQVPGKCCPEWVCDDGV--TPAQRSTAGHQLSALVTPASADAP 193  
Db 121 SEDVRLPSMDCPRRRTQVPGKCCPEWVCDDGV--TPAQRSTAGHQLSALVTPASADAP 180  
QY 194 CPNWSSTAMGECSTTCGIGIATRVSNQNRFCOLEIQRRLCLPRCLAASSHSSWNSAF 250  
Db 181 CPNWSSTAMGECSTTCGIGIATRVSNQNRFCOLEIQRRLCLSRPCLASRSHGSMNSAF 237

Search completed: July 25, 2003, 12:38:36  
Job time : 30 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 25, 2003, 12:35:05 ; Search time 23 Seconds  
(without alignments)  
1290.867 Million cell updates/sec

Title: US-10-010-408-2  
Perfect score: 1440  
Sequence: 1 MRGSPRLHLATSFLLISM.....LCLEPRCLAAARSHSSMNSAF 250

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 451899 seqs, 118759770 residues  
Total number of hits satisfying chosen parameters: 451899

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_AA:\*

1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCRT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCRTUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09D\_NEW\_PUB.pep:\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score  | Query Match | Length | ID               | Description       |
|------------|--------|-------------|--------|------------------|-------------------|
| 1          | 1440   | 100.0       | 250    | US-10-010-408-2  | Sequence 2, Appl  |
| 2          | 1323   | 91.9        | 227    | US-10-010-408-13 | Sequence 13, Appl |
| 3          | 1308.5 | 90.9        | 251    | US-10-112-267-20 | Sequence 20, Appl |
| 4          | 1303.5 | 90.5        | 250    | US-10-112-267-78 | Sequence 78, Appl |
| 5          | 1298.5 | 90.2        | 249    | US-10-112-267-79 | Sequence 79, Appl |
| 6          | 1297.5 | 89.8        | 248    | US-10-112-267-80 | Sequence 80, Appl |
| 7          | 1291.5 | 89.7        | 247    | US-10-112-267-81 | Sequence 81, Appl |
| 8          | 1284.5 | 89.2        | 246    | US-10-112-267-82 | Sequence 82, Appl |
| 9          | 1280.5 | 88.9        | 245    | US-10-112-267-83 | Sequence 83, Appl |
| 10         | 1276.5 | 88.6        | 244    | US-10-112-267-84 | Sequence 84, Appl |
| 11         | 1268.5 | 88.1        | 243    | US-10-112-267-85 | Sequence 85, Appl |
| 12         | 1264.5 | 87.8        | 242    | US-10-112-267-86 | Sequence 86, Appl |
| 13         | 1260.5 | 87.5        | 241    | US-10-112-267-87 | Sequence 87, Appl |
| 14         | 1257.5 | 87.3        | 239    | US-10-112-267-89 | Sequence 89, Appl |
| 15         | 1257.5 | 87.3        | 240    | US-10-112-267-88 | Sequence 88, Appl |

|    |        |      |     |                   |                    |
|----|--------|------|-----|-------------------|--------------------|
| 16 | 1253.5 | 87.0 | 238 | US-10-112-267-90  | Sequence 90, Appl  |
| 17 | 1247.5 | 86.6 | 237 | US-10-112-267-91  | Sequence 91, Appl  |
| 18 | 1243.5 | 86.4 | 236 | US-10-112-267-92  | Sequence 92, Appl  |
| 19 | 1234.5 | 85.7 | 235 | US-10-112-267-93  | Sequence 93, Appl  |
| 20 | 1232.5 | 85.6 | 234 | US-10-112-267-94  | Sequence 94, Appl  |
| 21 | 1228.5 | 85.3 | 233 | US-10-112-267-95  | Sequence 95, Appl  |
| 22 | 1224.5 | 85.0 | 232 | US-10-112-267-96  | Sequence 96, Appl  |
| 23 | 1219.5 | 84.7 | 231 | US-10-112-267-97  | Sequence 97, Appl  |
| 24 | 1217.5 | 84.5 | 229 | US-10-112-267-99  | Sequence 99, Appl  |
| 25 | 1217.5 | 84.5 | 230 | US-10-112-267-98  | Sequence 98, Appl  |
| 26 | 1216.5 | 84.5 | 228 | US-10-112-267-19  | Sequence 19, Appl  |
| 27 | 1064   | 73.9 | 250 | US-09-915-582-53  | Sequence 53, Appl  |
| 28 | 1064   | 73.9 | 250 | US-09-915-582-69  | Sequence 69, Appl  |
| 29 | 1064   | 73.9 | 250 | US-10-028-072-320 | Sequence 320, Appl |
| 30 | 1064   | 73.9 | 250 | US-10-121-049-320 | Sequence 320, Appl |
| 31 | 1064   | 73.9 | 250 | US-10-123-904-320 | Sequence 320, Appl |
| 32 | 1064   | 73.9 | 250 | US-10-140-470-320 | Sequence 320, Appl |
| 33 | 1064   | 73.9 | 250 | US-10-175-746-320 | Sequence 320, Appl |
| 34 | 1064   | 73.9 | 250 | US-10-176-918-320 | Sequence 320, Appl |
| 35 | 1064   | 73.9 | 250 | US-10-176-921-320 | Sequence 320, Appl |
| 36 | 1064   | 73.9 | 250 | US-10-137-865-320 | Sequence 320, Appl |
| 37 | 1064   | 73.9 | 250 | US-10-140-474-320 | Sequence 320, Appl |
| 38 | 1064   | 73.9 | 250 | US-10-142-431-320 | Sequence 320, Appl |
| 39 | 1064   | 73.9 | 250 | US-10-143-114-320 | Sequence 320, Appl |
| 40 | 1064   | 73.9 | 250 | US-10-140-002-320 | Sequence 320, Appl |
| 41 | 1064   | 73.9 | 250 | US-10-142-419-320 | Sequence 320, Appl |
| 42 | 1064   | 73.9 | 250 | US-10-123-262-320 | Sequence 320, Appl |
| 43 | 1064   | 73.9 | 250 | US-10-142-423-320 | Sequence 320, Appl |
| 44 | 1064   | 73.9 | 250 | US-10-121-050-320 | Sequence 320, Appl |
| 45 | 1064   | 73.9 | 250 | US-10-141-755-320 | Sequence 320, Appl |

## ALIGNMENTS

RESULT 1  
US-10-010-408-2  
Sequence 2, Application US/10010408  
Publication No. US20020165185A1  
GENERAL INFORMATION:  
APPLICANT: John J. Castelliot, Jr.  
TITLE OF INVENTION: No. US20020165185A1el Heparin-Induced CCN-Like Molecules  
and Uses Therefor  
NUMBER OF SEQUENCES: 13  
CORRESPONDING ADDRESSES:  
ADDRESSEE: LAHIVE & COCKFIELD, LLP  
STREET: 28 State Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/010,408  
FILING DATE: 07-Dec-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/044,273  
FILING DATE: March 19, 1998  
APPLICATION NUMBER: <Unknown>  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Amy E. Mandragouras  
REGISTRATION NUMBER: 36,207  
REFERENCE/DOCKET NUMBER: MBI-004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)742-4214  
INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:  
LENGTH: 250 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-10-010-408-2

Query Match 100.0%; Score 1440; DB 14; Length 250;  
Best Local Similarity 100.0%; Pred. No. 6,9e-183;  
Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRGSLHLATSFCLLSMCAQICRTPCTCPMPPPCGGCVPLVLDGCGCKVCARRL 60  
DB 1 MRGSLHLATSFCLLSMCAQICRTPCTCPMPPPCGGCVPLVLDGCGCKVCARRL 60  
QY 61 GSCDHLAVCDPSQGLVCPGAGPGHGAVALLEDSDSCVENGRYLDGETFKNCVYL 120  
DB 61 GSCDHLAVCDPSQGLVCPGAGPGHGAVALLEDSDSCVENGRYLDGETFKNCVYL 120  
QY 121 CRCDGFTCLPLCSEDEVRLPSMDCPRPKRIQVPGKCCPEWVCDOGVTPAIQSTAGCHQ 180  
DB 121 CRCDGFTCLPLCSEDEVRLPSMDCPRPKRIQVPGKCCPEWVCDOGVTPAIQSTAGCHQ 180  
QY 181 LSAVTPASADAPCPNMSTANGPCSTTGLGIATRVSNQNRFCOLEIORLCLPRPCLAA 240  
DB 181 LSAVTPASADAPCPNMSTANGPCSTTGLGIATRVSNQNRFCOLEIORLCLPRPCLAA 240  
QY 241 RSHSSMNSAF 250  
DB 241 RSHSSMNSAF 250

RESULT 2  
US-10-010-408-13  
Sequence 13, Application US/10010408  
Publication No. US20020165185A1  
GENERAL INFORMATION:  
APPLICANT: John J. Castellot, Jr.  
TITLE OF INVENTION: No. US20020165185A1 Heparin-Induced CCN-Like Molecules  
and Uses Therefor  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD, LLP  
STREET: 28 State Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/010,408  
FILING DATE: 07-Dec-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/044,273  
FILING DATE: March 19, 1998  
APPLICATION NUMBER: <Unknown>  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Amy E. Mandragouras  
REGISTRATION NUMBER: 36,207  
REFERENCE/DOCKET NUMBER: MB1-004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)742-4214  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 227 amino acids

TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 13:  
US-10-010-408-13

Query Match 91.9%; Score 1323; DB 14; Length 227;  
Best Local Similarity 100.0%; Pred. No. 3,8e-103;  
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 QLCRPTCPMPPPCGGCVPLVLDGCGCKVCARRLGEDSDHLVCDPSQGLVCPGAG 83  
DB 1 QLCRPTCPMPPPCGGCVPLVLDGCGCKVCARRLGEDSDHLVCDPSQGLVCPGAG 60  
QY 84 PGHGAVALLEDSDSCVENGRYLDGETFKNCVYLRCDCDGFCLPLCSEDEVRLPSW 143  
DB 61 PGHGAVALLEDSDSCVENGRYLDGETFKNCVYLRCDCDGFCLPLCSEDEVRLPSW 120  
QY 144 DCPKRIQVPGKCCPEWVCDOGVTPAIQSTAGCHQLSAVTPASADAPCPNMSTANGP 203  
DB 121 DCPKRIQVPGKCCPEWVCDOGVTPAIQSTAGCHQLSAVTPASADAPCPNMSTANGP 180  
QY 204 CSTTCGLGIATRVSNQNRFCOLEIORLCLPRPCLAAHSSMNSAF 250  
DB 181 CSTTCGLGIATRVSNQNRFCOLEIORLCLPRPCLAAHSSMNSAF 227

RESULT 3  
US-10-112-267-20  
Sequence 20, Application US/10112267  
Publication No. US20030068678A1  
GENERAL INFORMATION:  
APPLICANT: Botstein, David A.  
APPLICANT: Cohen, Robert  
APPLICANT: Goddard, Audrey  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: Lawrence, David A.  
APPLICANT: Levine, Arnold J.  
APPLICANT: Pennica, Diane  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME  
FILE REFERENCE: P1176r2  
CURRENT APPLICATION NUMBER: US/10/112,267  
PRIOR FILING DATE: 2002-03-27  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/182,145B  
PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-29  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/063,704  
PRIOR FILING DATE: EARLIER FILING DATE: 1997-10-29  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/073,612  
PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-04  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/081,695  
PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-14  
NUMBER OF SEQ ID NOS: 156  
SEQ ID NO 20  
LENGTH: 251  
TYPE: PRT  
ORGANISM: Mus musculus  
US-10-112-267-20

Query Match 90.9%; Score 1308.5; DB 15; Length 251;  
Best Local Similarity 90.0%; Pred. No. 6,8e-102;  
Matches 226; Conservative 9; Mismatches 15; Indels 1; Gaps 1;

QY 1 MRGSLHLATSFCLLSMCAQICRTPCTCPMPPPCGGCVPLVLDGCGCKVCARRL 60  
DB 1 MRGSLHLATSFCLLSMCAQICRTPCTCPMPPPCGGCVPLVLDGCGCKVCARRL 60  
QY 61 GSCDHLAVCDPSQGLVCPGAGPGHGAVALLEDSDSCVENGRYLDGETFKNCVYL 120  
DB 61 GSCDHLAVCDPSQGLVCPGAGPGHGAVALLEDSDSCVENGRYLDGETFKNCVYL 120





;; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME  
;; FILE REFERENCE: P1176R2  
;; CURRENT APPLICATION NUMBER: US/10/112,267  
;; PRIOR FILING DATE: 2002-03-27  
;; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/182,145B  
;; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-29  
;; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/063,704  
;; PRIOR FILING DATE: EARLIER FILING DATE: 1997-10-29  
;; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/073,612  
;; PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-04  
;; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/081,695  
;; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-14  
;; NUMBER OF SEQ ID NOS: 156  
;; SEQ ID NO 80  
;; LENGTH: 248  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-10-112-267-80

Query Match 89.8%; Score 1292.5; DB 15; Length 248;  
Best Local Similarity 89.9%; Pred. No. 1.5e-100;  
Matches 223; Conservative 9; Mismatches 15; Indels 1; Gaps 1;

QY 4 SPLHLATSFCLLSMWCAQLCRTCTCPMTPTPOCGVPLVLDGCGCCVCARRLGSS 63  
DB 1 LPLHLATSFCLLSMWVSQLCPAPACPMTPPOCGVPLVLDGCGCCVCARRLGSS 60  
QY 64 CDHLVCDPSQGLVCPGAGPGGHGAVCLLDEDDGSCENVGRYRLDGETFKPNCRYLCRC 123  
DB 61 CDHLVCDPSQGLVCPGAGPGGHGAVCLLDEDDGSCENVGRYRLDGETFKPNCRYLCRC 120  
QY 124 DGGFTCLPLCSEEDVRLPSWDCPRPRRIQVPGRCCEVWCDQAVMOPALOPSSAQCHQJLS 182  
DB 121 DGGFTCLPLCSEEDVRLPSWDCPRPRRIQVPGRCCEVWCDQAVMOPALOPSSAQCHQJLS 180  
QY 183 ALVTPASADAPCPNMSTAMGPGCTGCGIATRVSNONRFOCLEIORLCLPRLCLARS 242  
DB 181 ALVTPASADGCPNMSTAMGPGCTGCGIATRVSNONRFOCLEIORLCLPRLCLARS 240  
QY 243 HSSWNSAF 250  
DB 241 HSSWNSAF 248

RESULT 7  
US-10-112-267-81  
;; Sequence 81, Application US/10112267  
;; Publication No. US20030068678A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Bolstein, David A.  
;; APPLICANT: Cohen, Robert  
;; APPLICANT: Goddard, Audrey  
;; APPLICANT: Gurney, Austrey L.  
;; APPLICANT: Hillan, Kenneth J.  
;; APPLICANT: Lawrence, David A.  
;; APPLICANT: Levine, Arnold J.  
;; APPLICANT: Pennica, Diane  
;; APPLICANT: Roy, Margaret Ann  
;; APPLICANT: Wood, William I.  
;; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME  
;; FILE REFERENCE: P1176R2  
;; CURRENT APPLICATION NUMBER: US/10/112,267  
;; PRIOR FILING DATE: 2002-03-27  
;; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/182,145B  
;; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-29  
;; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/063,704  
;; PRIOR FILING DATE: EARLIER FILING DATE: 1997-10-29  
;; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/073,612  
;; PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-04  
;; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/081,695  
;; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-14  
;; NUMBER OF SEQ ID NOS: 156  
;; SEQ ID NO 81

;; LENGTH: 247  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-10-112-267-81

Query Match 89.7%; Score 1291.5; DB 15; Length 247;  
Best Local Similarity 90.3%; Pred. No. 1.8e-100;  
Matches 223; Conservative 8; Mismatches 15; Indels 1; Gaps 1;

QY 5 PLHLATSFCLLSMWCAQLCRTCTCPMTPTPOCGVPLVLDGCGCCVCARRLGSSC 64  
DB 1 PLHLATSFCLLSMWVSQLCPAPACPMTPPOCGVPLVLDGCGCCVCARRLGSSC 60  
QY 65 DHLVCDPSQGLVCPGAGPGGHGAVCLLDEDDGSCENVGRYRLDGETFKPNCRYLCRC 124  
DB 61 DHLVCDPSQGLVCPGAGPGGHGAVCLLDEDDGSCENVGRYRLDGETFKPNCRYLCRC 120  
QY 125 DGGFTCLPLCSEEDVRLPSWDCPRPRRIQVPGRCCEVWCDQAVMOPALOPSSAQCHQJLS 183  
DB 121 DGGFTCLPLCSEEDVRLPSWDCPRPRRIQVPGRCCEVWCDQAVMOPALOPSSAQCHQJLS 180  
QY 184 LVTTPASADAPCPNMSTAMGPGCTGCGIATRVSNONRFOCLEIORLCLPRLCLARS 243  
DB 181 LVTTPASADGCPNMSTAMGPGCTGCGIATRVSNONRFOCLEIORLCLPRLCLARS 240  
QY 244 SSWNSAF 250  
DB 241 SSWNSAF 247

RESULT 8  
US-10-112-267-82  
;; Sequence 82, Application US/10112267  
;; Publication No. US20030068678A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Bolstein, David A.  
;; APPLICANT: Cohen, Robert  
;; APPLICANT: Goddard, Audrey  
;; APPLICANT: Gurney, Austrey L.  
;; APPLICANT: Hillan, Kenneth J.  
;; APPLICANT: Lawrence, David A.  
;; APPLICANT: Levine, Arnold J.  
;; APPLICANT: Pennica, Diane  
;; APPLICANT: Roy, Margaret Ann  
;; APPLICANT: Wood, William I.  
;; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME  
;; FILE REFERENCE: P1176R2  
;; CURRENT APPLICATION NUMBER: US/10/112,267  
;; PRIOR FILING DATE: 2002-03-27  
;; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/182,145B  
;; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-29  
;; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/063,704  
;; PRIOR FILING DATE: EARLIER FILING DATE: 1997-10-29  
;; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/073,612  
;; PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-04  
;; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/081,695  
;; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-14  
;; NUMBER OF SEQ ID NOS: 156  
;; SEQ ID NO 82  
;; LENGTH: 246  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-10-112-267-82

Query Match 89.2%; Score 1284.5; DB 15; Length 246;  
Best Local Similarity 90.2%; Pred. No. 6.7e-100;  
Matches 223; Conservative 8; Mismatches 15; Indels 1; Gaps 1;

QY 6 LHLATSFCLLSMWCAQLCRTCTCPMTPTPOCGVPLVLDGCGCCVCARRLGSSCD 65  
DB 1 LHLATSFCLLSMWVSQLCPAPACPMTPPOCGVPLVLDGCGCCVCARRLGSSCD 60  
QY 66 HLHVCDPSQGLVCPGAGPGGHGAVCLLDEDDGSCENVGRYRLDGETFKPNCRYLCRC 125

```
DB 61 HLHVCDSQGLVCCPGAGPSGRGAVCLFEEDDGSCEVNGRRYLDGETFKPNCRYLCRCDD 120
QY 126 GGFCLPLCSEDEVRLPSMDCPRPRIQVPGKCCPEWVCDOGV-TPAIORSTAQHQHLSAL 184
DB 121 GGFCLPLCSEDEVRLPSMDCPRPRIQVPGKCCPEWVCDOAVMPAIOPSAOGHLSAL 180
QY 185 VTPASADAPCPNMSTANGPSTTCGLGIATRVSNQNRFCOLEIQRRLCLPRPCLARSHS 244
DB 181 VTPASADAPCPNMSTANGPSTTCGLGIATRVSNQNRFCOLEIQRRLCLSRCLASRSHG 240
QY 245 SWSNAF 250
DB 241 SWSNAF 246

RESULT 9.
US-10-112-267-83
; Sequence 83, Application US/10112267
; Publication No. US20030068678A1
; GENERAL INFORMATION:
; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/10/112,267
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/182,145B
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/063,704
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/073,612
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-04
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/081,695
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 83
; LENGTH: 245
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-112-267-83

Query Match 88.9%; Score 1280.5; DB 15; Length 245;
Best Local Similarity 90.2%; Pred. No. 1.4e-99;
Matches 221; Conservative 8; Mismatches 15; Indels 1; Gaps 1;
```

```
QY 7 HLHATSFLLISWYCAQLCTPCTCPWTPOCCPGVPLVLDGCGCCVCAARRLGECDH 66
DB 1 HLHATSFLLISWYCAQLCTPCTCPWTPOCCPGVPLVLDGCGCCVCAARRLGECDH 60
QY 67 LHVCDPSQGLVCCPGAGPSGRGAVCLFEEDDGSCEVNGRRYLDGETFKPNCRYLCRCDDG 126
DB 61 LHVCDPSQGLVCCPGAGPSGRGAVCLFEEDDGSCEVNGRRYLDGETFKPNCRYLCRCDDG 120
QY 127 GGFCLPLCSEDEVRLPSMDCPRPRIQVPGKCCPEWVCDOGV-TPAIORSTAQHQHLSALV 185
DB 121 GGFCLPLCSEDEVRLPSMDCPRPRIQVPGKCCPEWVCDOAVMPAIOPSAOGHLSALV 180
QY 186 TPAASADAPCPNMSTANGPSTTCGLGIATRVSNQNRFCOLEIQRRLCLPRPCLARSHS 245
DB 181 TPAASADAPCPNMSTANGPSTTCGLGIATRVSNQNRFCOLEIQRRLCLSRCLASRSHG 240
QY 246 WNSAF 250
DB 241 WNSAF 245
```

```
RESULT 10
US-10-112-267-84
; Sequence 84, Application US/10112267
; Publication No. US20030068678A1
; GENERAL INFORMATION:
; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/10/112,267
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/182,145B
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/063,704
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/073,612
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-04
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/081,695
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 84
; LENGTH: 244
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-112-267-84
```

```
Query Match 88.6%; Score 1276.5; DB 15; Length 244;
Best Local Similarity 90.2%; Pred. No. 3.1e-99;
Matches 220; Conservative 8; Mismatches 15; Indels 1; Gaps 1;
```

```
QY 8 HLHATSFLLISWYCAQLCTPCTCPWTPOCCPGVPLVLDGCGCCVCAARRLGECDH 67
DB 1 HLHATSFLLISWYCAQLCTPCTCPWTPOCCPGVPLVLDGCGCCVCAARRLGECDH 60
QY 68 HVCDSQGLVCCPGAGPSGRGAVCLFEEDDGSCEVNGRRYLDGETFKPNCRYLCRCDDG 127
DB 61 HVCDSQGLVCCPGAGPSGRGAVCLFEEDDGSCEVNGRRYLDGETFKPNCRYLCRCDDG 120
QY 128 GGFCLPLCSEDEVRLPSMDCPRPRIQVPGKCCPEWVCDOGV-TPAIORSTAQHQHLSALV 186
DB 121 GGFCLPLCSEDEVRLPSMDCPRPRIQVPGKCCPEWVCDOAVMPAIOPSAOGHLSALV 180
QY 187 PASADAPCPNMSTANGPSTTCGLGIATRVSNQNRFCOLEIQRRLCLPRPCLARSHS 246
DB 181 PASADAPCPNMSTANGPSTTCGLGIATRVSNQNRFCOLEIQRRLCLSRCLASRSHG 240
QY 247 NSAF 250
DB 241 NSAF 244

RESULT 11
US-10-112-267-85
; Sequence 85, Application US/10112267
; Publication No. US20030068678A1
; GENERAL INFORMATION:
; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
```



Db 1 AISPCLISWVYSQLCPACPCPMTPPQCPGVPVLVDGCGCCRCVACARLGECDHLHVC 60  
Qy 71 DPSGLVCPGAGGGHGAVALLEDDEDDGSCVNGRRYLDGETFKPNCRVLCRCDDGGFTTC 130  
Db 61 DPSGLVCPGAGGGHGAVALLEDDEDDGSCVNGRRYLDGETFKPNCRVLCRCDDGGFTTC 120  
Qy 131 LPLCSEDEVRLPSPWDCPRPRRIQVPGKCCPEWVCDQGV-TPAIORSTAGHQLSALVTPAS 189  
Db 121 LPLCSEDEVRLPSPWDCPRPRRIQVPGKCCPEWVCDQAVMOPAIQPSAQHQLSALVTPAS 180  
Qy 190 ADACCPMWTAMGPGCSTTCGIGATIRVSNONRFOLEIQRRLCLPRPCLAASSHSSWNSA 249  
Db 181 ADGCPMWTAMGPGCSTTCGIGATIRVSNONRFOLEIQRRLCLSRPCLAASSHSSWNSA 240  
Qy 250 F 250  
Db 241 F 241

RESULT 14  
; Sequence 89, Application US/10112267  
; Publication No. US20030068678A1  
; GENERAL INFORMATION:  
; APPLICANT: Botstein, David A.  
; APPLICANT: Cohen, Robert  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth J.  
; APPLICANT: Lawrence, David A.  
; APPLICANT: Levine, Arnold J.  
; APPLICANT: Pennica, Diane  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Wood, William I.  
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: P1176R2  
; CURRENT APPLICATION NUMBER: US/10/112,267  
; PRIOR FILING DATE: 2002-03-27  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/182,145B  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-29  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/063,704  
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-10-29  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/073,612  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-04  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/081,695  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-14  
; NUMBER OF SEQ ID NOS: 156  
; SEQ ID NO 89  
; LENGTH: 239  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-112-267-89

Query Match 87.3%; Score 1257.5; DB 15; Length 239;  
Best Local Similarity 90.4%; Pred. No. 1.2e-97;  
Matches 216; Conservative 8; Mismatches 14; Indels 1; Gaps 1;

Qy 13 SFLLLSWVCAQLCRTCTCPTWTPPQCPGVPVLVDGCGCCRCVACARLGECDHLHVC 72  
Db 1 SFLLLSWVCAQLCRTCTCPTWTPPQCPGVPVLVDGCGCCRCVACARLGECDHLHVC 60  
Qy 73 SGLVCPGAGPGGGHGAVALLEDDEDDGSCVNGRRYLDGETFKPNCRVLCRCDDGGFTCLP 132  
Db 61 SGLVCPGAGPGGGHGAVALLEDDEDDGSCVNGRRYLDGETFKPNCRVLCRCDDGGFTCLP 120  
Qy 133 LPLCSEDEVRLPSPWDCPRPRRIQVPGKCCPEWVCDQGV-TPAIORSTAGHQLSALVTPASAD 191  
Db 121 LPLCSEDEVRLPSPWDCPRPRRIQVPGKCCPEWVCDQAVMOPAIQPSAQHQLSALVTPASAD 180  
Qy 192 APCPMTAMGPGCSTTCGIGATIRVSNONRFOLEIQRRLCLPRPCLAASSHSSWNSAF 250  
Db 181 GPCPMTAMGPGCSTTCGIGATIRVSNONRFOLEIQRRLCLSRPCLAASSHSSWNSAF 239

RESULT 15  
; Sequence 88, Application US/10112267  
; Publication No. US20030068678A1  
; GENERAL INFORMATION:  
; APPLICANT: Botstein, David A.  
; APPLICANT: Cohen, Robert  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth J.  
; APPLICANT: Lawrence, David A.  
; APPLICANT: Levine, Arnold J.  
; APPLICANT: Pennica, Diane  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Wood, William I.  
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: P1176R2  
; CURRENT APPLICATION NUMBER: US/10/112,267  
; PRIOR FILING DATE: 2002-03-27  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/182,145B  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-29  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/063,704  
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-10-29  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/073,612  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-04  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/081,695  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-14  
; NUMBER OF SEQ ID NOS: 156  
; SEQ ID NO 88  
; LENGTH: 240  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-112-267-88

Query Match 87.3%; Score 1257.5; DB 15; Length 240;  
Best Local Similarity 90.4%; Pred. No. 1.2e-97;  
Matches 216; Conservative 8; Mismatches 14; Indels 1; Gaps 1;

Qy 13 SFLLLSWVCAQLCRTCTCPTWTPPQCPGVPVLVDGCGCCRCVACARLGECDHLHVC 72  
Db 2 SFLLLSWVCAQLCRTCTCPTWTPPQCPGVPVLVDGCGCCRCVACARLGECDHLHVC 61  
Qy 73 SGLVCPGAGPGGGHGAVALLEDDEDDGSCVNGRRYLDGETFKPNCRVLCRCDDGGFTCLP 132  
Db 62 SGLVCPGAGPGGGHGAVALLEDDEDDGSCVNGRRYLDGETFKPNCRVLCRCDDGGFTCLP 121  
Qy 133 LPLCSEDEVRLPSPWDCPRPRRIQVPGKCCPEWVCDQGV-TPAIORSTAGHQLSALVTPASAD 191  
Db 122 LPLCSEDEVRLPSPWDCPRPRRIQVPGKCCPEWVCDQAVMOPAIQPSAQHQLSALVTPASAD 181  
Qy 192 APCPMTAMGPGCSTTCGIGATIRVSNONRFOLEIQRRLCLPRPCLAASSHSSWNSAF 250  
Db 182 GPCPMTAMGPGCSTTCGIGATIRVSNONRFOLEIQRRLCLSRPCLAASSHSSWNSAF 240

Search completed: July 25, 2003, 12:39:05  
Job time : 24 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 25, 2003, 12:32:35 ; Search time 39 Seconds  
(without alignments)  
616.466 Million cell updates/sec

Title: US-10-010-408-2  
Perfect score: 1440  
Sequence: 1 MRCSPLHLATSLFLSLM.....LCLPRCLARSHSNNSAF 250

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : PIR\_76:.\*  
1: p1r1:.\*  
2: p1r2:.\*  
3: p1r3:.\*  
4: p1r4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID     | Description        |
|------------|-------|-------------|--------|----|--------|--------------------|
| 1          | 564   | 39.2        | 349    | 2  | A40551 | connective tissue  |
| 2          | 550.5 | 38.2        | 348    | 2  | A40578 | beta IG-M2 protein |
| 3          | 541.5 | 37.6        | 351    | 2  | S20078 | NOV protein - chic |
| 4          | 541   | 37.6        | 375    | 2  | A41428 | CEF-10 protein pre |
| 5          | 519   | 36.0        | 357    | 2  | I38069 | gene novh protein  |
| 6          | 500.5 | 34.8        | 379    | 2  | A35669 | gene CYR61 protein |
| 7          | 161   | 11.2        | 1620   | 2  | T27283 | hypothetical prote |
| 8          | 146   | 10.1        | 1172   | 1  | TSHU2  | hypothetical prote |
| 9          | 137.5 | 9.5         | 5376   | 2  | T42215 | zonalhesin - mouse |
| 10         | 136.5 | 9.5         | 1036   | 2  | T17405 | scavenger receptor |
| 11         | 136   | 9.4         | 1574   | 2  | T13954 | MEGF6 protein - ra |
| 12         | 135   | 9.4         | 2321   | 2  | S78549 | notch3 protein - h |
| 13         | 132.5 | 9.2         | 1111   | 2  | T26972 | hypothetical prote |
| 14         | 131.5 | 9.1         | 593    | 1  | GYHU   | granulin precursor |
| 15         | 131.5 | 9.1         | 1492   | 2  | A40333 | collagen alpha 1(I |
| 16         | 131.5 | 9.1         | 2910   | 2  | T42214 | collagen alpha 1(I |
| 17         | 130.5 | 9.1         | 1486   | 1  | B40333 | collagen alpha 1(I |
| 18         | 130   | 9.0         | 1034   | 2  | JCS598 | collagen alpha 1(I |
| 19         | 128.5 | 8.9         | 884    | 2  | T18649 | hypothetical prote |
| 20         | 128.5 | 8.9         | 1496   | 1  | CGH02V | collagen alpha 2(V |
| 21         | 128   | 8.9         | 1178   | 1  | A39804 | thrombospondin pre |
| 22         | 127   | 8.8         | 2531   | 2  | A46019 | notch-1 protein -  |
| 23         | 126.5 | 8.8         | 1466   | 1  | CGH07L | collagen alpha 1(I |
| 24         | 126   | 8.8         | 1251   | 2  | A57293 | latent transformin |
| 25         | 126   | 8.8         | 1964   | 2  | T09059 | notch4 - mouse     |
| 26         | 125.5 | 8.7         | 1487   | 1  | CGH06C | collagen alpha 1(I |
| 27         | 124.5 | 8.6         | 810    | 2  | T10756 | Net-homolog protei |
| 28         | 124   | 8.6         | 837    | 2  | A42112 | mucin-like peptide |
| 29         | 124   | 8.6         | 886    | 2  | I50694 | collagen alpha 1(I |

|    |       |     |      |   |        |                    |
|----|-------|-----|------|---|--------|--------------------|
| 30 | 124   | 8.6 | 1408 | 2 | S16148 | gene serrate prote |
| 31 | 124   | 8.6 | 1487 | 2 | B41182 | collagen alpha 1(I |
| 32 | 123   | 8.5 | 1444 | 2 | T18856 | angiogenesis inhib |
| 33 | 122.5 | 8.5 | 1188 | 2 | D86236 | protein F14N23.5 ( |
| 34 | 122.5 | 8.5 | 2201 | 2 | A32160 | tenascin-C - human |
| 35 | 122   | 8.5 | 304  | 2 | A32601 | insulin-like growt |
| 36 | 122   | 8.5 | 305  | 2 | I48601 | insulin-like growt |
| 37 | 122   | 8.5 | 305  | 2 | JW0508 | insulin-like growt |
| 38 | 122   | 8.5 | 589  | 2 | B38128 | epithelin/granulin |
| 39 | 121.5 | 8.4 | 282  | 2 | S50031 | prostacyclin-stimu |
| 40 | 121.5 | 8.4 | 591  | 2 | I48141 | acroganin - guine  |
| 41 | 121.5 | 8.4 | 2476 | 2 | T34022 | zonalhesin - pig   |
| 42 | 120   | 8.3 | 317  | 2 | I46916 | insulin-like growt |
| 43 | 120   | 8.3 | 615  | 2 | A05269 | collagen alpha 1(I |
| 44 | 120   | 8.3 | 1023 | 2 | T30257 | IGF rc binding pro |
| 45 | 120   | 8.3 | 2813 | 1 | VWU    | von Willebrand fac |

## ALIGNMENTS

RESULT 1  
A40551  
Connective tissue growth factor - human  
C:Species: Homo sapiens (man)  
C:Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 21-Jul-2000  
C:Accession: A40551; S44205  
R:Bradham, D.M.; Igarashi, A.; Potter, R.L.; Grotenhorst, G.R.  
J. Cell Biol. 114, 1285-1294, 1991  
A:Title: Connective tissue growth factor: a cysteine-rich mitogen secreted by human v  
A:Reference number: A40551; MUID:91373462; PMID:1654338  
A:Accession: A40551  
A:Molecule type: mRNA  
A:Residues: 1-349 <BR>  
A:Cross-references: GB:M92934; GB:M36965; GB:S56201; NID:G180923; PIDN:AAA91279.1; PI  
R:Oemar, B.S.; Werner, A.; Yang, Z.; Garner, J.M.; Gentz, R.; Luescher, T.F.  
submitted to the EMBL Data Library, April 1994  
A:Description: Differential cloning and expression of human connective tissue growth  
A:Reference number: S44205  
A:Accession: S44205  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-349 <OEM>  
A:Cross-references: EMBL:X78947; NID:9474933; PID:9474934  
Query Match 39.2%, Score 564, DB 2, Length 349;  
Best Local Similarity 46.9%, Pred. No. 2.1e-35;  
Matches 112; Conservative 28; Mismatches 81; Indels 18; Gaps 6;  
Qy 13 SFLCLSM---VCAQLCRPFCPCPTMP-PQCPGVPPLVDGCGCKVCARRLGESCDHL 67  
Db 12 AFVVELACSRPAVAGQMGSPCRCPDEPAPCPAGVSLVLDGCGCCCAQDLBELCER 71  
Qy 68 HVCPSQGLVQCPAGFGHGAVALDLDDSGCEVNGHRYLDGETFPKRCVFLCRCDGG 127  
Db 72 DPCDPHKGFLFCDFSPANRKTIGVCTA-KDGAPEJFGGVYVSGSFSOSCKYQCTCLDGA 130  
Qy 128 FTCLPUCSEVRLPSMDCPRPKRIQVPEKCCPEWVCGDGVTPAIQIRSTAGHQSAL--- 184  
Db 131 VGMPLCSMDVRLSPDCPPFRKRLPKCCSEEWCDCE---PKDQ---TVGPPALAAARLE 185  
Qy 185 ---VTPASADAPCPNNSWTGPGCTGGLIATRVSNQNFQLEIQRICLPRPCA 239  
Db 186 DTFGPDPPIATNANCLVOTTEMSACSKTGMSITRVYNDANASRLKQSRICLMPCEA 244

RESULT 2  
A40578  
beta IG-M2 protein precursor - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 06-Mar-1992 #sequence\_revision 06-Mar-1992 #text\_change 01-Dec-2000  
C:Accession: A40578; A53228  
R:Brunner, A.; Chinn, J.; Neubaer, M.; Purchio, A.F.





Query Match 36.0%; Score 519; DB 2; Length 357;  
 Best Local Similarity 42.3%; Pred. No. 5.2e-32;  
 Matches 102; Conservative 32; Mismatches 85; Indels 22; Gaps 7;

OY 6 LIHLALSFLCLLSMVCACQLCRTPC--TCPTPPCCPGCVPLVLDGCGCCYCARRLGES 63  
 DB 22 LHLHLG-----GVAATQRCPCPCPGRCPPATPCAGVRAVLDGSCCCVCAQRGES 74  
 OY 64 CDHLHVCDSQGLVCGQPGAGPGHGVAVCLLDEDDGSCSEVNGRRYLDGTFKPNCRVLCRC 123  
 DB 75 CSDLPCDESSGLTCDNSADPNSNOTGCTAVEGD-NCVFDEVTIRSGEKFPSCFQCTC 133  
 OY 124 DDGFTCLPLCSEDEVRLPSPMCPRPKRIOVPGKCCPEWVC--DQ-----GVTPAIQRSTA 176  
 DB 134 RDGIGCVPRCQDLVLPENPCPAAPRAKVEPGECEKIKGCPDEEDLSGLTLAYRREA 193  
 OY 177 OGHQSLALVTPASADAPCPNMSTWGPCSTTCGIGIATRVSNQNFQLEIQRRLCPRP 236  
 DB 194 ---TLGVEVSDSVN--CIEQTTWTACSKSCGMSGSTRVTRNRNOCHEMLKQTRLCVVRP 248  
 OY 237 C 237  
 DB 249 C 249

RESULT 6

A35669  
 gene CYR61 protein precursor - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 28-Sep-1990 #sequence\_revision 18-Nov-1992 #text\_change 05-Nov-1999  
 C:Accession: A35669; I48319; S16446  
 R:O'Brien, T.P.; Yang, G.P.; Sanders, L.; Lau, L.F.  
 Mol. Cell. Biol. 10, 3569-3577, 1990  
 A:Title: Expression of cyr61, a growth factor-inducible immediate-early gene.  
 A:Reference number: A35669; MUID:90287146; PMID:2353916  
 A:Accession: A35669  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-379 <RES>  
 A:Cross-references: GB:M32400; NID:q192909; PIDN:AAA37512.1; PID:q309206  
 A>Note: the authors translated the codon GAT for residue 337 as Gln  
 R:LaLinkic, B.V.; O'Brien, T.P.; Lau, L.F.  
 Nucleic Acids Res. 19, 3261-3267, 1991  
 A:Title: Promoter function and structure of the growth factor-inducible immediate early  
 A:Reference number: I48319; MUID:91288203; PMID:2062642  
 A:Accession: I48319  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-379 <RES>  
 A:Cross-references: EMBL:X56790; NID:q50632; PIDN:CAA40109.1; PID:q50633  
 A>Note: the authors did not translate the codon for residue 108  
 A:Note: the authors translated the codon GAT for residue 337 as Gln  
 C:Genetics:  
 A:Gene: CYR61  
 A:Introns: 21/3; 93/1; 208/1; 279/3  
 C:Superfamily: von Willebrand factor type C repeat homology  
 F:99-166/Domain: von Willebrand factor type C repeat homology <VMC>

Query Match 34.8%; Score 500.5; DB 2; Length 379;  
 Best Local Similarity 35.6%; Pred. No. 1.4e-30;  
 Matches 96; Conservative 35; Mismatches 106; Indels 33; Gaps 2;  
 OY 1 MRGSPHLHLATSFCLLSWCAQLCRTPCPTPPCCPGCVPLVLDGCGCCYCARRL 60  
 DB 1 MSSSTFRLAVAVLHLHLTRALSTCPAACHCPLEAPCAPGVGLVRGCGCCYCAKQL 60  
 OY 61 GESCDHLHVCDSQGLVCGQPGAGPGHGVAVCLLDEDDGSCSEVNGRRYLDGTFKPNCRV 120  
 DB 61 NEDCSKQPCDHTGLEGNEFGASSALKGICRAOSEGRPCYNSRIYNGSFQPCNCHQ 120  
 OY 121 CRCDDGGTCLPLCSEDEVRLPSPMCPRPKRIOVPGKCCPEWVCQD----- 165  
 DB 121 CTCIDGAVGCIPLCPOELISGHCSCQPCNRLVAVKVSQCCFEMVCDSDSIKDSLDDDDLLG 180

OY 166 -----GVTPAIQRSTAQGHQSLALVTPASADA-PCPNMSTWGPCSTT 207  
 DB 181 IDASVEITRNNELIAIKGSSSLKRLPYFGIEPRVLPFLAHAGKCIYQTTSMSCSKS 240

OY 208 CGLGATRVSNQNFQLEIQRRLCPRP 237  
 DB 241 CGTGISTRVNTNDNEPCRLVKEFTRICEVAPC 270

RESULT 7

hypothetical protein Y64G10A.f - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 C:Accession: T27283  
 R:Ainscough, R.  
 submitted to the EMBL Data Library, September 1999  
 A:Reference number: Z20336  
 A:Accession: T27283  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-1620 <WIL>  
 A:Cross-references: EMBL:AL110498; NID:e1542303; PIDN:CAB54471.1; CESP:Y64G10A.f  
 A:Experimental source: clone Y64G10A  
 C:Genetics:  
 A:Gene: CESP:Y64G10A.f  
 A:Introns: 77/1; 116/1; 198/1; 282/1; 365/1; 425/1; 466/1; 548/1; 559/1; 601/1; 625/1

Query Match 11.2%; Score 161; DB 2; Length 1620;  
 Best Local Similarity 26.3%; Pred. No. 0.00015;  
 Matches 59; Conservative 18; Mismatches 93; Indels 54; Gaps 11;

OY 30 CTCPP--WTPPCCPGCVPLVLDGCGCCYCARRLGESCDHL-HVCDPSSGLV---CQPGAG 83  
 DB 1124 CRCPAGWTGPDCCQTSPLGRHGEGRHSCQCSNAGSCDRVTGFCDCPSGFGKNCESRCP 1183  
 OY 84 PGHGAVCL-----LDEDDGSCSEVNGRRYLDGFT-----FRPNCRYLCRCD 124  
 DB 1184 EGLWGSNCKMKKCLCMHGGECKENGDCEC-----IDGWTGSLCPFGFGNCRNCRCK 1238  
 OY 125 DGGFTCLPLCSEDEVRLPSPMCPRPKRIOVPG---KCCPEWVCDOGVTPAIQRSTAQGHQ 180  
 DB 1239 NGA-SCDRKTCRCCLCPWMSGEHCESCVSGHYKACKECTCECENG----- 1283  
 OY 181 LSAIVTPASADAPC-PNMSTWGPCSTTCGIGIATRVSNQNFRC 223  
 DB 1284 --ALCDPISGHCSCQPGWRGK--KCNRPCLKGYGRHCSQSCRC 1323

RESULT 8

Thrombospondin 2 precursor - human  
 C:Species: Homo sapiens (man)  
 C:Date: 19-May-1995 #sequence\_revision 03-Aug-1995 #text\_change 13-Aug-1999  
 C:Accession: A47379; A42173  
 R:Labell, T.L.; Milewicz, D.J.; Distche, C.M.; Byers, P.H.  
 Genomics 17, 225-229, 1993  
 A:Title: Sequence and characterization of the complete human thrombospondin 2 cDNA: p  
 A:Reference number: A47379; MUID:94010892; PMID:8406456  
 A:Accession: A47379  
 A:Molecule type: mRNA  
 A:Residues: 1-1172 <LAB>  
 A:Cross-references: GB:L12350; NID:q307505; PIDN:AAA03703.1; PID:q307506  
 R:Labell, T.L.; Milewicz, D.J.; Distche, C.M.; Byers, P.H.  
 Genomics 12, 421-429, 1992  
 A:Title: Thrombospondin II: partial cDNA sequence, chromosome location, and expressio  
 A:Reference number: A42173; MUID:92217961; PMID:1559694  
 A:Accession: A42173  
 A:Molecule type: mRNA  
 A:Residues: 560-1172 <LAB>  
 A:Cross-references: GB:M81339  
 A:Experimental source: fibroblast





C:Accession: JCI284; A38128; A38118; A36698; B36698; G36698; D36698; A56873  
 R:Bhandari, V.; Bateman, A.  
 Biochem. Biophys. Res. Commun. 188, 57-63, 1992  
 A:Title: Structure and chromosomal location of the human granulin gene.  
 A:Reference number: JCI284; MUID:93038704; PMID:1417868  
 A:Accession: JCI284  
 A:Molecule type: DNA  
 A:Residues: 1-593 <BNA>  
 R:Plowman, G.D.; Green, J.M.; Neubauer, M.G.; Buckley, S.D.; McDonald, V.L.; Todaro, G.J.  
 J. Biol. Chem. 267, 13073-13078, 1992  
 A:Title: The epithelin precursor encodes two proteins with opposing activities on epithelial cells.  
 A:Reference number: A38128; MUID:92317004; PMID:1618805  
 A:Accession: A38128  
 A:Status: Preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-593 <PLD>  
 A:Cross-references: GB:X62320; NID:931192; PIDN:CAAA4196.1; PID:931193  
 R:Bhandari, V.; Palfrey, R.G.E.; Bateman, A.  
 Proc. Natl. Acad. Sci. U.S.A. 89, 1715-1719, 1992  
 A:Title: Isolation and sequence of the granulin precursor cDNA from human bone marrow.  
 A:Reference number: A38118; MUID:92179253; PMID:1542665  
 A:Accession: A38118  
 A:Molecule type: mRNA  
 A:Residues: 1-406, 'R', 408-433, 'G', 435-453, 'G', 455-459, 'O', 461-546, 'A', 548-566, 'R', 568-593  
 A:Cross-references: GB:W5161; NID:9183612; PIDN:AA58617.1; PID:9183613  
 A:Note: This sequence has been revised in reference JCI284  
 R:Bateman, A.; Belcourt, D.; Bennett, H.; Lazure, C.; Solomon, S.  
 Biochem. Biophys. Res. Commun. 173, 1161-1168, 1990  
 A:Title: Granulins, a novel class of peptide from leukocytes.  
 A:Reference number: A36698; MUID:91097544; PMID:2266320  
 A:Accession: A36698  
 A:Molecule type: protein  
 A:Residues: 281-336 <BAT>  
 A:Note: this protein was purified and characterized as granulin A  
 A:Accession: B36698  
 A:Molecule type: protein  
 A:Residues: 206-218, 'H', 220-233 <BA2>  
 A:Note: this protein was purified and characterized as granulin B  
 A:Accession: C36698  
 A:Molecule type: protein  
 A:Residues: 364-367, 'X', 369-385, 'H', 387-396 <BA3>  
 A:Note: this protein was purified and characterized as granulin C  
 A:Accession: D36698  
 A:Molecule type: protein  
 A:Residues: 442-446, 'X', 456-458, 'DG' <BA4>  
 R:Kardana, A.; Bagshaw, K.D.; Coles, B.; Read, D.; Taylor, M.  
 Br. J. Cancer 67, 686-692, 1993  
 A:Title: Characterisation of UGP and its relationship with beta-core fragment.  
 A:Reference number: A56873; MUID:9329246; PMID:8471426  
 A:Accession: A56873  
 A:Molecule type: protein  
 A:Residues: 281-283, 'X', 285-289, 'S', 291-295 <KAR>  
 A:Experimental source: urine  
 A:Note: sequence extracted from NCBI backbone (NCBI:129524)  
 C:Genetics:  
 A:Gene: GDB:GRN  
 A:Cross-references: GDB:136006; OMIM:138945  
 A:Map position: 17pter-17qter  
 A:Introns: 46/3; 88/3; 117/1; 154/3; 200/1; 236/3; 279/1; 311/3; 393/3; 471/3; 548/3  
 C:Superfamily: granulin  
 C:Keywords: glycoprotein; tandem repeat  
 F:1-17/Domain: signal sequence #status predicted <SIG>  
 F:18-593/Product: granulin #status predicted <MAV>  
 F:18-593/Product: progranulin #status predicted <PRO>  
 F:18-44/Product: paraganulin #status experimental <PGR>  
 F:58-113/Product: granulin G #status predicted <GRG>  
 F:123-119/Product: granulin F #status predicted <GRF>  
 F:206-261/Product: granulin B #status experimental <GRB>  
 F:281-336/Product: granulin A #status experimental <GRA>  
 F:364-417/Product: granulin C #status experimental <GRC>  
 F:442-496/Product: granulin E #status predicted <GRE>  
 F:518-573/Product: granulin D #status predicted <GRD>  
 F:366/Binding site: carbohydrate (asn) (covalent) #status experimental

Query Match 9.1%; Score 131.5; DB 1; Length 593;  
 Best Local Similarity 24.6%; Pred. No. 0.011;  
 Matches 62; Conservative 15; Mismatches 78; Indels 97; Gaps 16;

|    |     |  |
|----|-----|--|
| QY | 32  | CPWTPPCPOGVPVLVDCGCGCKV-----CARRLESC--DHLAVCDPSGGLVCPQPG 83  |
| DB | 208 | CDPARSACV-----DSTCCLEPSGKYGCCPMENATCCSHLCC--PDPTVCDL--- 255  |
| QY | 84  | PGHGAVCLDDE-----DGSEVNGRRYLDGEFTFKPNCNRVLCRCDDGDF 128        |
| DB | 256 | ---IOSKLSKENATDILLTKLPANTVGDYK-----DMEVSCPDEYTCRLQSGAW 304   |
| QY | 129 | TCLP-----LCSEVRLPMSWDCPRPKIOPVPCCP-----EWVCDQGYTPAIOGSTA 176 |
| DB | 305 | GCCEFTQAVCCED-----HIH-----CCPAGFTCDTKQKCTCGDGPQHWMEKA 348    |
| QY | 177 | OGHLSALVTPAS--ADAPCPMWSFAMGPGSTTGLGATRVSNONRRCOLEIQRRLCP 234 |
| DB | 349 | PAH--LSLPDPQALRDPVCDNVSSC--PSSDTC-----COLTSGEMGCCP 390       |
| QY | 235 | RP---CLAARSH 243   |
| DB | 391 | IPAVCCSDHQH 402  |

RESULT 15  
 collagen alpha 1(II) chain precursor - African clawed frog  
 C:Species: Xenopus laevis (African clawed frog)  
 C:Date: 16-Sep-1992 #sequence\_revision 16-Sep-1992 #text\_change 16-Jul-1999  
 C:Accession: A40333  
 R:Su, M.W.; Suzuki, H.R.; Bleker, J.J.; Solursh, M.; Ramirez, F.  
 J. Cell Biol. 115, 565-575, 1991  
 A:Title: Expression of two nonallelic type II procollagen genes during Xenopus laevis  
 A:Reference number: A40333; MUID:92011898; PMID:1918153  
 A:Accession: A40333  
 A:Molecule type: mRNA  
 A:Status: nucleic acid sequence not shown  
 A:Cross-references: GB:M63596  
 A:Note: this sequence is presented as substitutions relative to another sequence in a  
 es they replace; the appropriate interpretation of the sequence figure was reconstruct  
 C:Keywords: collagen alpha 1(II) chain; fibrillar collagen carboxyl-terminal homolo  
 C:Keywords: coiled coil; extracellular matrix; glycoprotein; trimer; triple helix  
 F:37-96/Domain: von Willibrand factor type C repeat homology <VMC>  
 F:1263-1492/Domain: fibrillar collagen carboxyl-terminal homology <FCC>

Query Match 9.1%; Score 131.5; DB 2; Length 1492;  
 Best Local Similarity 36.7%; Pred. No. 0.024;  
 Matches 33; Conservative 12; Mismatches 26; Indels 19; Gaps 5;

|    |     |   |
|----|-----|---|
| QY | 94  | DEDD-----GSCENVRRYLDGEFTFKPNCNRVLCRCDDGFTCLPCLSEDPRLPMSDCPPRK 149 |
| DB | 28  | DEEDVLDGSCYQDQGRSDKQVMPKPCQICVDTG-----TVLDELICESSKDCP--- 80       |
| QY | 150 | RIQVP-GRKCEPMWCDQGVTPAIOGSTAG 178                                 |
| DB | 81  | NAEIRFEGCCP-----ICPLEQSTSSG 103                                   |

Search completed: July 25, 2003, 12:38:00  
 Job time : 41 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 25, 2003, 12:25:50 ; Search time 23 Seconds

(without alignments)  
511.160 Million cell updates/sec

Title: US-10-010-408-2

Perfect score: 1440

Sequence: 1 MRGSPILHLATSPFLCLSM.....LCLEPCLAAHSHSSMSNAF 250

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score  | Query Match | Length | ID | Description        |
|------------|--------|-------------|--------|----|--------------------|
| 1          | 1390   | 96.5        | 250    | 1  | CTGFL_RAT          |
| 2          | 1308.5 | 90.9        | 251    | 1  | Q9JHC6_rattus norv |
| 3          | 1064   | 73.9        | 250    | 1  | CTGFL_HUMAN        |
| 4          | 568    | 39.4        | 349    | 1  | CTGFL_BOVIN        |
| 5          | 564    | 39.2        | 349    | 1  | CTGFL_HUMAN        |
| 6          | 550.5  | 38.2        | 347    | 1  | CTGFL_RAT          |
| 7          | 550.5  | 38.2        | 346    | 1  | CTGFL_MOUSE        |
| 8          | 541.5  | 37.6        | 351    | 1  | NOV_CHICK          |
| 9          | 541.5  | 37.6        | 375    | 1  | CE10_CHICK         |
| 10         | 533    | 37.0        | 349    | 1  | CTGFL_PIG          |
| 11         | 532.5  | 37.0        | 353    | 1  | NOV_COTUA          |
| 12         | 529.5  | 36.8        | 354    | 1  | NOV_MOUSE          |
| 13         | 519    | 36.0        | 357    | 1  | NOV_HUMAN          |
| 14         | 516    | 35.8        | 351    | 1  | NOV_RAT            |
| 15         | 512    | 35.6        | 343    | 1  | NOV_XENLA          |
| 16         | 504    | 35.0        | 379    | 1  | CYR6_RAT           |
| 17         | 500.5  | 34.8        | 379    | 1  | CYR6_MOUSE         |
| 18         | 498.5  | 34.6        | 381    | 1  | CYR6_HUMAN         |
| 19         | 146    | 10.1        | 1172   | 1  | TSPP2_HUMAN        |
| 20         | 146    | 9.9         | 2282   | 1  | ZAN_RABIT          |
| 21         | 142    | 9.7         | 184    | 1  | ESM1_HUMAN         |
| 22         | 138    | 9.6         | 810    | 1  | NEL1_HUMAN         |
| 23         | 137.5  | 9.5         | 5376   | 1  | ZAN_MOUSE          |
| 24         | 135    | 9.4         | 2321   | 1  | NTRC3_HUMAN        |
| 25         | 133    | 9.2         | 830    | 1  | SREB_HUMAN         |
| 26         | 131.5  | 9.1         | 593    | 1  | GRN_HUMAN          |
| 27         | 130    | 9.0         | 867    | 1  | SSPO_BOVIN         |
| 28         | 130    | 9.0         | 2482   | 1  | VWF_PIG            |
| 29         | 128.5  | 8.9         | 281    | 1  | IBP7_MOUSE         |
| 30         | 128.5  | 8.9         | 1459   | 1  | CA12_MOUSE         |
| 31         | 128.5  | 8.9         | 1496   | 1  | CA25_HUMAN         |
| 32         | 128    | 8.9         | 1178   | 1  | TSPP2_CHICK        |
| 33         | 127.5  | 8.9         | 2812   | 1  | ZAN_HUMAN          |

|    |       |     |      |   |             |                     |
|----|-------|-----|------|---|-------------|---------------------|
| 34 | 127   | 8.8 | 2531 | 1 | NTRC1_MOUSE | Q01705 mus musculus |
| 35 | 126.5 | 8.8 | 1466 | 1 | CA13_HUMAN  | P02461 homo sapien  |
| 36 | 126   | 8.8 | 1964 | 1 | NTCA_MOUSE  | P31695 mus musculus |
| 37 | 125.5 | 8.7 | 588  | 1 | GRN_RAT     | P23785 r granulin   |
| 38 | 125   | 8.7 | 1629 | 1 | AT9_HUMAN   | Q9P244 homo sapien  |
| 39 | 124.5 | 8.6 | 810  | 1 | NEL1_RAT    | Q62919 rattus norv  |
| 40 | 124   | 8.6 | 837  | 1 | MOCL_RAT    | P98089 rattus norv  |
| 41 | 124   | 8.6 | 1262 | 1 | CA13_CHICK  | P12105 gallus gall  |
| 42 | 124   | 8.6 | 1408 | 1 | SRRR_DROME  | P18168 drosophila   |
| 43 | 123   | 8.5 | 2319 | 1 | NTRC3_RAT   | Q91772 rattus norv  |
| 44 | 122.5 | 8.5 | 2201 | 1 | TENA_HUMAN  | P24821 homo sapien  |
| 45 | 122.5 | 8.5 | 2911 | 1 | FBN2_HUMAN  | P35556 homo sapien  |

## ALIGNMENTS

```

RESULT 1
CTGFL_RAT
AC 09JHC6:
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Connective tissue growth factor-like protein precursor (CTGF-L) (WNT1
DE inducible signaling pathway protein 2) (WISP-2) (CCN family protein
DE COP-1)
GN WISP2 OR CTGFL OR COP1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98414629; PubMed=9742130;
RA Zhang R., Averboukh L., Zhu W., Zhang H., Jo H., Dempsey P.J.,
RA Coffey R.J., Pardoe A.B., Liang P.;
RT "Identification of rCop-1, a new member of the CCN protein family, as
RT a negative regulator for cell transformation.";
RL Mol. Cell. Biol. 18:6131-6141(1998).
CC
CC -1- FUNCTION: May play an important role in modulating bone turnover.
CC Promotes the adhesion of osteoblast cells and inhibits the binding
CC of fibrinogen to integrin receptors. In addition, inhibits
CC osteocalcin production (By similarity).
CC -1- SUBCELLULAR LOCATION: Secreted (Probable).
CC -1- SIMILARITY: Contains 1 IGFBP domain.
CC -1- SIMILARITY: Contains 1 WIF domain.
CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: AF259981; AAF69011.1;
CC InterPro: IPR000867; Insl_gro_fac_pr.
CC InterPro: IPR000884; TSP1.
CC InterPro: IPR001007; WIF_C.
CC Pfam: PF00219; IGFBP; 1.
CC Pfam: PF00090; TSP1; 1.
CC Pfam: PF00093; WIF; 1.
CC SMART: SM00121; IB; 1.
CC SMART: SM00209; TSP1; 1.
CC PROSITE: PS00222; IGF_BINDING; 1.
CC PROSITE: PS00092; TSP1; 1.
CC PROSITE: PS01208; WIF_C; 1.
CC PROSITE: PS0184; WIF_C2; 1.
CC Signal.
CC SIGNAL 1 23 POTENTIAL.

```

```

FT CHAIN 24 250 CONNECTIVE TISSUE GROWTH FACTOR-LIKE
FT DOMAIN 26 96 PROTEIN.
FT DOMAIN 98 164 IGFBP.
FT DOMAIN 194 238 WMFC.
FT CARBOHYD 196 196 TSP TYPE-1.
SO SEQUENCE 250 AA: 27005 MW: 9A147074628BC447 CRC64;
Query Match 96.5%; Score 1390; DB 1; Length 250;
Best Local Similarity 96.8%; Pred. No. 2.3e-98;
Matches 242; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 MRSPPIHLATSFILLSVCAQLCRPTCTCPMTPOGCVPLVLDGCGCCVCARL 60
DB 1 MRSPPIRLATSFILLSVCAQLCRPTCTCPMTPOGCVPLVLDGCGCCVCARL 60
QY 61 GESCDHLHVCDSOGIVCPGAGPGGAGVCLDEDDGSCVNGRRYLDGETFKPCRYL 120
DB 61 TESCDHLHVCDSOGIVCPGAGPGGAGVCLDEDDGSCVNGRRYLDGETFKPCRYL 120
QY 121 CRCDDGFTCLPLCSSEVRLPSWDCPRPRIOVPKCCPEWVCDOGVTPAIOGSTAQGH 180
DB 121 CRCDDGFTCLPLCSSEVRLPSWDCPRPRIOVPKCCPEWVCDOGVTPAIOGSTAQGH 180
QY 181 LSLALVPASADAPCPMSTAMGPCSTTCGLGATRVSNQRFQLEIQRRLCPCLAA 240
DB 181 LSLALVPASADAPCPMSTAMGPCSTTCGLGATRVSNQRFQLEIQRRLCPCLAA 240
QY 241 RSHSSWNSAF 250
DB 241 RSHSSWNSAF 250

RESULT 2
CTGL_MOUSE STANDARD: PRT: 251 AA.
ID CTGL_MOUSE 092064;
AC 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Connective tissue growth factor-like protein precursor (CTGF-L)
DE (WNT1 Inducible signaling pathway protein 2) (WISP-2).
DE WISP2 OR CTGFL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 11
RP SEQUENCE FROM N.A.
RC TISSUE-Mammary gland;
RX MEDLINE=99061933; PubMed=9843955;
RA Pennica D., Swanson T.A., Welsh J.W., Roy M.A., Lawrence D.A., Lee J.,
RA Brush J., Taneyhill L.A., Denel B., Lew M., Matanabe C., Cohen R.L.,
RA Melham M.F., Finley G.G., Quirke P., Goddard A.D., Hillan K.J.,
RA Gurney A.L., Botstein D., Levine A.J.;
RT "WISP genes are members of the connective tissue growth factor family
RT that are up-regulated in wnt-1-transformed cells and aberrantly
RT expressed in human colon tumors."
RT Proc. Natl. Acad. Sci. U.S.A. 95:14717-14722(1998).
RN 12
RP SEQUENCE FROM N.A.
RC TISSUE-Lung;
RX MEDLINE=99287915; PubMed=10358067;
RA Kumar S., Hand A.T., Connor J.R., Dodds R.A., Ryan P.J., Trill J.J.,
RA Fisher S.M., Nuttall M.E., Lipschutz D.B., Zou C., Hwang S.M.,
RA Votta B.J., James I.E., Rleman D.J., Gowen M., Lee J.C.;
RT "Identification and cloning of a connective tissue growth factor-like
RT cDNA from human osteoblasts encoding a novel regulator of osteoblast
RT functions."
RL J. Biol. Chem. 274:17123-17131(1999).
CC -1- FUNCTION: May play an important role in modulating bone turnover.
CC Promotes the adhesion of osteoblast cells and inhibits the binding
CC of fibrinogen to integrin receptors. In addition, inhibits

```

```

CC osteocalcin production (By similarity).
CC -1- SUBCELLULAR LOCATION: Secreted (Probable).
CC -1- SIMILARITY: Contains 1 IGFBP domain.
CC -1- SIMILARITY: Contains 1 WMFC domain.
CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
CC -----
DR EMBL: AF100778; AAC96320.1; -.
DR EMBL: AF126063; AAD18058.1; -.
DR MGI: MGI:1328326; Wisp2.
DR InterPro: IPR000867; Ins1_gro_fac_dr.
DR InterPro: IPR000884; TSP1.
DR InterPro: IPR001007; WMFC.
DR Pfam: PF00219; IGFBP; 1.
DR Pfam: PF00090; tsp.1; 1.
DR Pfam: PF00093; vwc; 1.
DR SMART: SM00121; TB; 1.
DR SMART: SM00209; TSP1; 1.
DR SMART: SM00214; VWC; 1.
DR PROSITE: PS00222; IGF_BINDING; 1.
DR PROSITE: PS50092; TSP1; 1.
DR PROSITE: PS01208; WMFC_1; 1.
DR PROSITE: PS50184; WMFC_2; 1.
DR Signal.
FT SIGNAL 1 23
FT CHAIN 24 251
FT FT 24 251
FT FT 26 96
FT FT 98 164
FT FT 194 238
FT FT 196 196
FT CARBOHYD 197 197
SO SEQUENCE 251 AA: 27095 MW: 893E1633FEB5C7FC CRC64;
Query Match 90.9%; Score 1308.5; DB 1; Length 251;
Best Local Similarity 90.0%; Pred. No. 3.2e-92;
Matches 226; Conservative 9; Mismatches 15; Indels 1; Gaps 1;

QY 1 MRSPPIHLATSFILLSVCAQLCRPTCTCPMTPOGCVPLVLDGCGCCVCARL 60
DB 1 MRSPPIRLATSFILLSVCAQLCRPTCTCPMTPOGCVPLVLDGCGCCVCARL 60
QY 61 GESCDHLHVCDSOGIVCPGAGPGGAGVCLDEDDGSCVNGRRYLDGETFKPCRYL 120
DB 61 GESCDHLHVCDSOGIVCPGAGPGGAGVCLDEDDGSCVNGRRYLDGETFKPCRYL 120
QY 121 CRCDDGFTCLPLCSSEVRLPSWDCPRPRIOVPKCCPEWVCDOGVTPAIOGSTAQGH 179
DB 121 CRCDDGFTCLPLCSSEVRLPSWDCPRPRIOVPKCCPEWVCDOGVTPAIOGSTAQGH 180
QY 181 LSLALVPASADAPCPMSTAMGPCSTTCGLGATRVSNQRFQLEIQRRLCPCLAA 239
DB 181 LSLALVPASADAPCPMSTAMGPCSTTCGLGATRVSNQRFQLEIQRRLCPCLAA 240
QY 240 ARSHSSWNSAF 250
DB 241 RSHSSWNSAF 251

RESULT 3
CTGL_HUMAN STANDARD: PRT: 250 AA.
ID CTGL_HUMAN 076076;
AC 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 15-SEP-2003 (Rel. 42, Last annotation update)
DE Connective tissue growth factor-like protein precursor (CTGF-L) (WNT1

```

DE Inducible signaling pathway protein 2) (WISP-2) (Connective tissue  
DE growth factor-related protein 58).  
GN WISP2 OR CTGF OR CT58 OR CCN5.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99287915; PubMed=10358067;  
RA Kumar S., Hand A.T., Connor J.R., Dodds R.A., Ryan P.J., Trill J.J.,  
RA Fisher S.M., Nuttall M.E., Lipschultz D.B., Zou C., Hwang S.M.,  
RA Voita B.J., James I.E., Riegan D.J., Gowen M., Lee J.C.;  
RT "Identification and cloning of a connective tissue growth factor-like  
RT cDNA from human osteoblasts encoding a novel regulator of osteoblast  
RT functions";  
RL J. Biol. Chem. 274:17123-17131(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99061933; PubMed=9843955;  
RA Pennica D., Swanson T.A., Welsh J.W., Roy M.A., Lawrence D.A., Lee J.,  
RA Brush J., Taneyhill L.A., Denel B., Lew M., Watanabe C., Cohen R.L.,  
RA Melham M.F., Finley G.G., Quirke P., Goddard A.D., Hillan K.J.,  
RA Gunney A.L., Botstein D., Levine A.J.;  
RT "WSP genes are members of the connective tissue growth factor family  
RT that are up-regulated in vnt-1-transformed cells and aberrantly  
RT expressed in human colon tumors";  
RL Proc. Natl. Acad. Sci. U.S.A. 95:14717-14722(1998).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21638749; PubMed=11780052;  
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,  
RA Jones M., Stavriles G., Almeida J.P., Babbage A.K., Baggaley C.L.,  
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,  
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,  
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,  
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,  
RA Clegg S., Copley V.E., Collier R.E., Connor R.E., Corby N.R.,  
RA Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,  
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,  
RA Gammam D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,  
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
RA Lehaegels M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,  
RA Marsh V.L., Martin S.L., McConachle L.J., McElay K., McMurtry A.A.,  
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,  
RA Oliver K., Parker A., Patel R., Pearce T.A.W., Peck A.I.,  
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,  
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Shownkeen R., Sims S.,  
RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,  
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
RA Tracey A.C., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,  
RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,  
RA Wilmshurst L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
RA Rogers J.;  
RT "The DNA sequence and comparative analysis of human chromosome 20.";  
RL Nature 414:865-871(2001).  
RN [5]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.R.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Pringle C.,  
RA Raha S.S., Logucliano N.A., Peters G.J., Abramson R.D., Mulhaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huliy S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,  
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
CC -1- FUNCTION: May play an important role in modulating bone turnover.  
CC Promotes the adhesion of osteoblast cells and inhibits the binding  
CC of fibrinogen to integrin receptors. In addition, inhibits  
CC osteocalcin production.  
CC -1- SUBCELLULAR LOCATION: Secreted (Probable).  
CC -1- TISSUE SPECIFICITY: Expressed in primary osteoblasts, fibroblasts,  
CC ovary, testes, and heart.  
CC -1- SIMILARITY: Contains 1 IGFBP domain.  
CC -1- SIMILARITY: Contains 1 WFC domain.  
CC -1- SIMILARITY: Contains 1 TSP type-1 domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: AF083500; AAC70350.1; -  
DR EMBL: AF074604; AAC26794.1; -  
DR EMBL: AF100780; AAC96322.1; -  
DR EMBL: AL139352; CAB94788.1; -  
DR EMBL: BC017782; AAH17782.1; -  
DR Genew: HGNC:12770; WISP2.  
DR MIM: 603399; -  
DR GO: GO:0005625; C:soluble fraction; TAS.  
DR GO: GO:0007267; P:cell-cell signaling; TAS.  
DR GO: GO:0007048; P:oncogenesis; TAS.  
DR GO: GO:0007165; P:signal transduction; TAS.  
DR InterPro: IPR000867; Insl\_gro\_fac\_pr.  
DR InterPro: IPR001007; WSP\_C.  
DR Pfam: PF00219; IGFBP; 1.  
DR Pfam: PF00090; tsp\_1; 1.  
DR Pfam: PF00093; wfc; 1.  
DR SMART: SM00121; IB; 1.  
DR SMART: SM00209; TSP1; 1.  
DR SMART: SM00214; WVC; 1.  
DR PROSITE: PS00222; IGF BINDING; 1.  
DR PROSITE: PS50092; TSP1; 1.  
DR PROSITE: PS01208; WFC\_1; 1.  
DR PROSITE: PS50184; WFC\_2; 1.  
KW SIGNAL.  
FT CHAIN 1 23 POTENTIAL.  
FT CHAIN 24 250 CONNECTIVE TISSUE GROWTH FACTOR-LIKE  
FT PROTEIN.  
FT DOMAIN 26 96 IGFBP.  
FT DOMAIN 98 164 WFC.  
FT DOMAIN 194 238 TSP TYPE-1.  
SQ SEQUENCE 250 AA; 26825 MM; C499837EF42FEAC CRC64;  
Query Match 73.98; Score 1064; DB 1; Length 250;  
Best Local Similarity 73.68; Pred. No. 9.2e-74;  
Matches 184; Conservative 17; Mismatches 49; Indels 0; Gaps 0;  
QY 1 MRGSPILHLATSLCLLSMWCAOLCRTPCPCWTPQCGVPLVLDGCGCCKVCARRL 60







GN CTGF.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20145935; PubMed=10679821;  
 RA Xu J., Smock S.L., Safadi F.F., Rosenzweig A.B., Odgren P.R.,  
 RA Marks S.C., Jr., Owen T.A., Popoff S.N.;  
 RT "Cloning the full-length cDNA for rat connective tissue growth factor:  
 RT implications for skeletal development.";  
 RL J. Cell. Biochem. 77:103-115(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Tezuka K., Tamatani T.;  
 RT "Rattus norvegicus connective tissue growth factor.";  
 RL Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.  
 CC -1- FUNCTION: MAJOR CONNECTIVE TISSUE MITOATTRACTANT SECRETED BY  
 CC VASCULAR ENDOTHELIAL CELLS. PROMOTES PROLIFERATION AND  
 CC DIFFERENTIATION OF CHONDROCYTES (BY SIMILARITY). MEDIATES CELL  
 CC ADHESION AND ENHANCES FIBROBLAST GROWTH FACTOR-INDUCED DNA  
 CC SYNTHESIS (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- SIMILARITY: Contains 1 IGFBP domain.  
 CC -1- SIMILARITY: Contains 1 TSP type-1 domain.  
 CC -1- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; AF120275; AAD39132.1; -;  
 DR EMBL; AB023068; BAA82125.1; -;  
 DR InterPro: IPR006208; Cys\_knot.  
 DR InterPro: IPR006207; Cys\_knot.  
 DR InterPro: IPR000867; Inbl\_gro\_fac-pr.  
 DR InterPro: IPR000864; TSP1.  
 DR Pfam: PF00219; IGFBP; 1.  
 DR Pfam: PF00090; TSP\_1; 1.  
 DR Pfam: PF00093; VWC; 1.  
 DR SMART; SM00041; CT; 1.  
 DR SMART; SM00121; IB; 1.  
 DR SMART; SM00209; TSP1; 1.  
 DR PROSITE: PS01185; VWC; 1.  
 DR PROSITE: PS01225; CTCK\_1; 1.  
 DR PROSITE: PS00222; IGFBP; 1.  
 DR PROSITE: PS50092; TSP1; 1.  
 DR PROSITE: PS01208; VWC\_2; 1.  
 DR PROSITE: PS01208; VWC\_2; 1.  
 KW Cell adhesion; DNA synthesis; Extracellular matrix; Signal.  
 FT SIGNAL 1 24  
 FT CHAIN 25 347  
 FT DOMAIN 31 98  
 FT DOMAIN 99 165  
 FT DOMAIN 166 241  
 FT DOMAIN 242 328  
 FT DISULFID 254 328  
 FT DISULFID 271 305  
 FT DISULFID 282 321  
 FT DISULFID 285 333  
 FT DISULFID 290 337  
 FT CONFLICT 35 35  
 FT CONFLICT 94 94  
 SQ SEQUENCE 347 AA: 37756 MW: 37756 MW: CFEELAI9766B7B16 CRC64;

Query Match 38.2%; Score 550.5; DB 1; Length 347;  
 Best Local Similarity 45.4%; Pred. No. 6.9e-35;  
 Matches 108; Conservative 28; Mismatches 83; Indels 19; Gaps 5;  
 QY 15 LCLLSMVC-----AQLCTPCTC-PWTPPOCPQGYPLVLDGCGCCVCARRLGESCDELH 68  
 DB 11 LALVILLCTRPATGDCSNOCCOAAEAPRCAGVSLVLDGCGCCVCARRLGESCDELH 70  
 QY 69 VDDPSQGLVCGPAGPBGHGAVALCLEDDEDCSEVNGRRRLDGEFPKPCRVLCRDDPGF 128  
 DB 71 PCDPFGKGLFCDGFSANRKGICVCTA-KDCAFCVFGGSYRSESEFQSSCKYCTCLDCAV 129  
 QY 129 TCLPLCSEEDVRLPSMDCPRKRIQVPGKCPPEWCDQGVTPAIRSTAQHQLSAL---- 184  
 DB 130 GCVPLCSMDVRLPSPDCPPRRVKLPKCCCEWVDEP-----KRTVVGPAALAYRLDE 184  
 QY 185 ---VTPASADAPCPWMSWAMGCSSTGIGIATRYNSNORFQGLIQRRLCPRLCA 239  
 DB 185 TFGPPTWIRANCLVQTTEWSACSTCGWISTRTVNTNTRKESRLCWVRPCEA 242  
 RESULT 7  
 CTGF\_MOUSE  
 ID CTGF\_MOUSE STANDARD; PRT; 348 AA.  
 AC P29268; Q92200;  
 DT 01-DEC-1992 (Rel. 24, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Connective tissue growth factor precursor (fisp-12 protein)  
 GN (Hypertrophic chondrocyte-specific protein 24).  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91363290; PubMed=1888698;  
 RA Ryseck R.-P., Macdonald-Bravo H., Mattei M.-G., Bravo R.;  
 RT "Structure, mapping, and expression of fisp-12, a growth factor-  
 RT inducible gene encoding a secreted cysteine-rich protein.";  
 RL Cell Growth Differ. 2:225-233(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91229639; PubMed=2029337;  
 RA Brunner A., Chinn J., Neubauer M.G., Purchio A.F.;  
 RT "Identification of a gene family regulated by transforming growth  
 RT factor-beta.";  
 RL DNA Cell Biol. 10:293-300(1991).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strusberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Scheffer G.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Ditschenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Frange C.J.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Boulfard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butcherfield Y.S.N., Krzyzinski M.I., Skalski U., Smallus D.E.,  
 RA Scherch A., Schin J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RN [4]  
 RP FUNCTION, AND SUBCELLULAR LOCATION.  
 RX MEDLINE-97327410; PubMed-9184077;  
 RA Kireeva M.L., Latick B.V., Kolesnikova T.V., Chen C.C., Yang G.P.,  
 RA Abler A.S., Lau L.F.;  
 RT "Cyt61 and Flsp12 are both ECM-associated signalling molecules;  
 RT activities, metabolism, and localization during development.";  
 RL Exp. Cell Res. 233:63-77(1997).  
 RN [5]  
 RP FUNCTION.  
 RX MEDLINE-99182484; PubMed-10082563;  
 RA Babic A.M., Chen C.C., Lau L.F.;  
 RT "Flsp12/mouse connective tissue growth factor mediates endothelial  
 RT cell adhesion and migration through integrin alphavbeta3, promotes  
 RT endothelial cell survival, and induces angiogenesis in vivo.";  
 RL Mol. Cell. Biol. 19:2958-2966(1999).  
 CC -1- FUNCTION: MAJOR CONNECTIVE TISSUE MITOATTRACTANT SECRETED BY  
 CC VASCULAR ENDOTHELIAL CELLS. PROMOTES PROLIFERATION AND  
 CC DIFFERENTIATION OF CHONDROCYTES (BY SIMILARITY). MEDIATES CELL  
 CC ADHESION AND ENHANCES FIBROBLAST GROWTH FACTOR-INDUCED DNA  
 CC SYNTHESIS (BY SIMILARITY).  
 CC -1- SUBUNIT: Monomer (By similarity).  
 CC -1- SUBCELLULAR LOCATION: FOUND IN THE EXTRACELLULAR MATRIX AND AS A  
 CC SOLUBLE FORM.  
 CC -1- TISSUE SPECIFICITY: TESTIS, SPLEEN, KIDNEY, LUNG, HEART, AND BRAIN  
 CC (LOWEST LEVEL IN TESTIS AND HIGHEST IN LUNG).  
 CC -1- INDUCTION: By growth factors.  
 CC -1- SIMILARITY: Contains 1 IGFBP domain.  
 CC -1- SIMILARITY: Contains 1 WFPC domain.  
 CC -1- SIMILARITY: Contains 1 TSP type-1 domain.  
 CC -1- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 CC -----  
 DR EMBL: M70641; AAA37627.1; -;  
 DR EMBL: M70642; AAA37628.1; -;  
 DR EMBL: M80263; AAA73335.1; -;  
 DR EMBL: BC006783; AAH06783.1; -;  
 DR PIR: A40578; A40578.  
 DR MGD: MGI:95537; Ctgf.  
 DR GO: GO:0005578; C:extracellular matrix; IDA.  
 DR GO: GO:0008201; F:heparin binding activity; IDA.  
 DR GO: GO:0005178; F:integrin binding activity; IDA.  
 DR GO: GO:0001525; P:angiogenesis; IDA.  
 DR GO: GO:0016477; P:cell migration; IDA.  
 DR GO: GO:0007160; P:cell-matrix adhesion; IDA.  
 DR GO: GO:0008543; P:FGF receptor signaling pathway; IDA.  
 DR GO: GO:0007229; P:integrin-mediated signaling pathway; IDA.  
 DR GO: GO:0001503; P:ossification; IMP.  
 DR InterPro: IPR006208; Cys\_knot.  
 DR InterPro: IPR006207; Cys\_knot\_C.  
 DR InterPro: IPR000867; Ins1\_gro\_fac\_pr.  
 DR InterPro: IPR001007; WFPC.  
 DR Pfam: PF00007; Cys\_knot; 1.  
 DR Pfam: PF00219; IGFBP; 1.  
 DR Pfam: PF00090; TSP\_1; 1.  
 DR Pfam: PF00093; vwc; 1.  
 DR SMART: SM00041; CT; 1.  
 DR SMART: SM00121; IB; 1.  
 DR SMART: SM00209; TSP1; 1.  
 DR SMART: SM00214; vwc; 1.  
 DR PROSITE: PS01185; CTCK\_1; 1.  
 DR PROSITE: PS01225; CTCK\_2; 1.  
 DR PROSITE: PS00222; IGF\_BINDING; 1.  
 DR PROSITE: PS00992; TSP1; 1.  
 DR PROSITE: PS01208; WFPC\_1; 1.

DR PROSITE: PS0184; WFPC\_2; 1.  
 KW Cell adhesion; DNA synthesis; Extracellular matrix; Signal.  
 FT SIGNAL 1 25 POTENTIAL.  
 FT CHAIN 26 348 CONNECTIVE TISSUE GROWTH FACTOR.  
 FT DOMAIN 32 99 IGFBP.  
 FT DOMAIN 100 166 WFPC..  
 FT DOMAIN 197 242 TSP TYPE-1.  
 FT DOMAIN 255 329 CTCK.  
 FT DISULFID 255 292 BY SIMILARITY.  
 FT DISULFID 272 306 BY SIMILARITY.  
 FT DISULFID 283 322 BY SIMILARITY.  
 FT DISULFID 286 324 BY SIMILARITY.  
 FT DISULFID 291 328 BY SIMILARITY.  
 FT CONFLICT 161 161 E -> K (IN REF. 1).  
 SQ SEQUENCE 348 AA; 37794 MW; 4D7B6D9089174049 CRC64;  
 Query Match 38.2%; Score 550.5; DB 1; Length 348;  
 Best local Similarity 45.6%; Pred. No. 6.9e-35;  
 Matches 109; Conservative 26; Mismatches 81; Indels 23; Gaps 5;  
 QY 18 LSMVCAQLCRTP-----CTC-PWTPPCPGVPVLYLDGCCCKYCARLGSCDHL 67  
 Db 11 LAIVLALCTRPATGQCSACQCAEAARHPAGVSLVLDGCCCKRYCAQLDELCTER 70  
 QY 68 HVCPSQGLVCPAGPCGCGAVCLLDEDDSCVNGRRYLDGTFKPNCRVLCRDGCG 127  
 Db 71 DPCDPHKLFCDFCSPPANKRIGVCTA-KDGAFCVFGSGVRSGESFOSSCKYQCTCLDGA 129  
 QY 128 FTCLPLCESEVRLSPMDCPKPKRIQVPGCKCEVWCDGVPALQIRSAQGHQSLAL--- 184  
 Db 130 VGVPLCGMDVRLSPDCPPFRVRYKLPRKCCSEWCDP-----KDRVAPALAAVRL 184  
 QY 185 ----VTPASADAPCPNNSTANGPCSTTGCLGIATRVSNQNRFCOLEIQRRLCPRIA 239  
 Db 185 DTFGPDPTPMRANLVQVTEHSASCKTGMSIRVTNDNFCLKQSRCLCMVRPCEA 243  
 RESULT 8  
 NOV\_CHICK  
 ID NOV\_CHICK STANDARD; PRT; 351 AA.  
 AC P28686;  
 DT 01-DEC-1992 (Rel. 24, Created)  
 DT 01-DEC-1992 (Rel. 24, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE NOV protein precursor (Nephroblastoma overexpressed gene protein).  
 GN NOV.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OC NCBI\_TaxID=9031;  
 RP [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-Brown leghorn;  
 RX MEDLINE-92107157; PubMed-1309586;  
 RA Joliet V., Marthnerie C., Dambrine G., Plassiart G., Brisac M.,  
 RA Crochet J., Perbal B.;  
 RT "Proviral rearrangements and overexpression of a new cellular gene  
 RT (nov) in myeloblastosis-associated virus type 1-induced  
 RT nephroblastomas.";  
 RL Mol. Cell. Biol. 12:10-21(1992).  
 CC -1- FUNCTION: IMMEDIATE-EARLY PROTEIN LIKELY TO PLAY A ROLE IN CELL  
 CC GROWTH REGULATION. ITS OVEREXPRESSION IS ASSOCIATED WITH  
 CC TUMORIGENESIS AND EXPRESSION OF A N-TERMINAL-TRUNCATED VERSION  
 CC OF NOV GENE IN CHICKEN EMBRYONIC FIBROBLASTS (CEF) IS SUFFICIENT  
 CC TO INDUCE THE TRANSFORMATION OF CEF IN VITRO.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- TISSUE SPECIFICITY: BRAIN AND HEART, AND AT A LOWER LEVEL IN  
 CC MUSCLE AND INTESTINE, IN THE EMBRYO. LUNG AND LESS SO IN BRAIN AND  
 CC SPLEEN, IN ADULT CHICKEN.  
 CC -1- DEVELOPMENTAL STAGE: NAVI-INDUCED NEPHROBLASTOMAS EXPRESS A HIGH  
 CC LEVEL OF NOV GENE WHOSE TRANSCRIPTION IS NORMALLY ARRESTED IN  
 CC ADULT KIDNEY.



```

QY 10 LATSFLCLSMWCAOLCRPTCTPWTPOCPGQVPLVLDGCGCCYKCARLGECDHLHV 69
DB 10 LAALALCLARLALSGPCRAVCCGPAAPQCAAGVGLVVDGCGCCYKCARLGECDHLHV 69
QY 70 CDPGGLVQCPGAGFGHGVCLLDDGSCFVNGRRYLDGETFPKRCRVLCRDDGGFT 129
DB 70 CDHTKGLSCNFGASPAATNGICRAQSEGRPEYKXIVNGESFQPNCKHQCTCIDGAVG 129
QY 130 CLPLCSEVDRLPSMCPKPRKIQVGRKCCPEWVCCGV-----TPAIDRS 174
DB 130 CLPLCSEVDRLPSMCPKPRKIQVGRKCCPEWVCCGV-----TPAIDRS 174
QY 175 TAQCHQSLAV-----TPASADAP-----CPNWTAMGPGSTCGLGIAITRVSNO 219
DB 190 LTRNNELAIYVAGKMLPVGSESPQSAFENPKRIVQTTSMQSKTCGISTIRVIND 249
QY 220 NRCQLEIQRRLCLPRPC 237
DB 250 NDCRLIKETRICEVPRC 267

RESULT 10
CTGF_PIG STANDARD: PRT: 349 AA.
AC 019113;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
GN Connective tissue growth factor precursor.
OS Sus. scrofa (Pig).
OC Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Eumetazoa; Eutheria; Cetiartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=uterus;
RX MEDLINE=97390475; PubMed=9242708;
RA Harding P.A.; Steffen C.L., Kim G.Y., Vegunta R.K., Diehl J.R.,
RT "Purification and characterization of novel heparin-binding growth
RT factors in uterine secretory fluids. Identification as heparin-
RT regulated Mr 10,000 forms of connective tissue growth factor.";
RL J. Biol. Chem. 272:20275-20282(1997).
CC -1- FUNCTION: MAJOR CONNECTIVE TISSUE MITOATTRACTANT SECRETED BY
CC VASCULAR ENDOTHELIAL CELLS. PROMOTES PROLIFERATION AND
CC ADHESION AND ENHANCES FIBROBLAST GROWTH FACTOR-INDUCED DNA
CC SYNTHESIS (BY SIMILARITY).
CC -1- SUBUNIT: Monomer (By similarity).
CC -1- SUBCELLULAR LOCATION: FOUND IN THE EXTRACELLULAR MATRIX AND AS A
CC SOLUBLE FORM (BY SIMILARITY).
CC -1- SIMILARITY: Contains 1 IGFBP domain.
CC -1- SIMILARITY: Contains 1 WFPC domain.
CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
CC -1- SIMILARITY: Contains 1 C-terminal cystine knot-like (CTCK) domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: 083916; AAC48756.1; .
CC InterPro: IPR006208; Cys_knot.
CC InterPro: IPR006207; Cys_knot.C.
CC InterPro: IPR000867; Insl_gro_fac.pr.
CC InterPro: IPR000884; TSP1.
CC InterPro: IPR001007; WF_C.
CC Pfam: PF00007; Cys_knot; 1.

```

```

DR Pfam: PF00219; IGFBP; 1.
DR Pfam: PF00090; tsp.1; 1.
DR Pfam: PF00093; vwc; 1.
DR SMART: SM00041; CT; 1.
DR SMART: SM00121; TB; 1.
DR SMART: SM00209; TSP1; 1.
DR SMART: SM00214; vwc; 1.
DR PROSITE: PS01185; CTCK.1; 1.
DR PROSITE: PS01225; CTCK.2; 1.
DR PROSITE: PS00222; IGF_BINDING; 1.
DR PROSITE: PS50092; TSP1; 1.
DR PROSITE: PS01208; WF_C.1; 1.
DR PROSITE: PS50184; WF_C.2; 1.
KW Cell adhesion; DNA synthesis; Extracellular matrix; signal.
FT SIGNAL 1 26
FT CHAIN 1 349
FT DOMAIN 33 100
FT DOMAIN 101 167
FT DOMAIN 198 243
FT DOMAIN 256 330
FT DISULFID 256 293
FT DISULFID 273 307
FT DISULFID 284 323
FT DISULFID 287 325
FT DISULFID 292 329
SQ SEQUENCE 349 AA; 38007 MW; BB510E2B52D4A0 CRC64;

Query Match 37.0%; Score 533; DB 1; Length 349;
Best Local Similarity 44.4%; Pred. No. 1.5e-33;
Matches 106; Conservative 30; Mismatches 85; Indels 18; Gaps 5;

QY 13 SFCLLSM-----VCAOLCRPTCTPWTPO-CPGQVPLVLDGCGCCYKCARLGECDHL 67
DB 12 AFVLLALCLSPRASQDCGCGCAGARRACPAVSILVLDGCGCRLCAQLGELCTER 71
QY 68 HVCDFPSQGLVCPGAGPGHGVCLLDDGSCFVNGRRYLDGETFPKRCRVLCRDDGG 127
DB 72 DPCDHPKGLFDFGSPARKRIGVCA-KDGAFCVGVGVYNSGESFQSSCKKQCCLDGA 130
QY 128 FTCLPLCSEVDRLPSMCPKPRKIQVGRKCCPEWVCCGVTPAIDRSTAGHQSAL--- 184
DB 131 VGVCLPLCSEVDRLPSMCPKPRKIQVGRKCCPEWVCCGVTPAIDRSTAGHQSAL--- 184
QY 185 ----TPASADAPCPNWTAMGPGSTCGLGIAITRVSNONRRCOLEIQRRLCLPRPCA 239
DB 186 DTFGDPPTMARANCLVQTTSMQSKTCGISTIRVINDNAFCRLKESRLCAQVPRCEA 244

RESULT 11
NOV_COTUA STANDARD: PRT: 353 AA.
AC P42642;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
GN NOV protein precursor (Nephroblastoma overexpressed gene protein).
OS Coturnix coturnix japonica (Japanese quail).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauromorpha; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Coturnix.
OX NCBI_TaxID=93934;
RN [1]
RP SEQUENCE FROM N.A.
RA Weiskirchen R., Bister K.;
RL Submitted (AUG-1994) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: IMMEDIATE-EARLY PROTEIN LIKELY TO PLAY A ROLE IN CELL
CC GROWTH REGULATION (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -1- SIMILARITY: Contains 1 IGFBP domain.
CC -1- SIMILARITY: Contains 1 WFPC domain.
CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
CC -1- SIMILARITY: Contains 1 C-terminal cystine knot-like (CTCK) domain.

```

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>  
 CC or send an email to [license@isb-sdb.ch](mailto:license@isb-sdb.ch)).  
 CC  
 DR EMBL: U13063; AAA2128.1; -  
 DR InterPro: IPR006208; Cys\_knot.  
 DR InterPro: IPR006207; Cys\_knot\_C.  
 DR InterPro: IPR000867; Insl\_gro\_fac-pr.  
 DR InterPro: IPR000884; TSP1.  
 DR InterPro: IPR001007; VWF\_C.  
 DR Pfam: PF00007; Cys\_knot; 1.  
 DR Pfam: PF00219; IGFBP; 1.  
 DR Pfam: PF00090; TSP\_1; 1.  
 DR Pfam: PF00093; VWC; 1.  
 DR SMART: SM00041; CT; 1.  
 DR SMART: SM00121; IB; 1.  
 DR SMART: SM00209; TSP1; 1.  
 DR SMART: SM0214; VWC; 1.  
 DR PROSITE: PS01185; CTCK\_1; 1.  
 DR PROSITE: PS01225; CTCK\_2; 1.  
 DR PROSITE: PS00222; IGF\_BINDING; 1.  
 DR PROSITE: PS00922; TSP1; 1.  
 DR PROSITE: PS01208; VWF\_C\_1; 1.  
 DR PROSITE: PS0184; VWF\_C\_2; 1.  
 KW Proto-oncogene; Growth factor; Signal.  
 FT SIGNAL 1 26  
 FT CHAIN 1 26  
 FT DOMAIN 27 353  
 FT DOMAIN 33 105  
 FT DOMAIN 106 172  
 FT DOMAIN 203 248  
 FT DOMAIN 260 334  
 FT DISULFID 260 297  
 FT DISULFID 277 311  
 FT DISULFID 288 327  
 FT DISULFID 291 329  
 FT DISULFID 296 333  
 FT CARBOHYD 276 276  
 FT SEQUENCE 353 AA; 38667 MW; 717D9F85338028B96 CRC64; .)  
 Query Match 37.0%; Score 532.5; DB 1; Length 353;  
 Best Local Similarity 45.7%; Pred. No. 1.6e-33;  
 Matches 101; Conservative 32; Mismatches 73; Indels 15; Gaps 5;  
 26 CRTPC--TCPTPTPCPCPGVPLVLDGCGCCVCARLGGSCDHLVHCDPSGGLVQCPAG 83  
 33 CRRPCGGRCAPAPRCAPVPAVLDDGCGCLVCAQRGSCSPLLPCDSDSGGLYCDRGE 92  
 84 PGHGAVCLLDLDDSGCEVNGRRYLDGETFKPCRYLCRCDGFTCLPCLCSEDEVLPFM 143  
 93 DGGGIGICVLEGD--NCVFDMIRNGEFPQSCYKQCGRCRGQICLPRLNGLILPFP 151  
 144 DCPFRKRIQVPGKCCPEWVCDQ-----GVPAIORSAGHQLSALVTPASADAPCPN 196  
 152 DCPFRKRIQVPGKCCPEWVCDQ-----GVPAIORSAGHQLSALVTPASADAPCPN 196  
 153 DCPFRKRIQVPGKCCPEWVCDQ-----GVPAIORSAGHQLSALVTPASADAPCPN 196  
 197 WSTANGPCSTGGLGATRVSNONRFOLEIQRILCPPC 237  
 207 QTTEMSACSRSCGMGSHVTRVNRNOCCPEWVKTQRLCKMPPC 247

DE gene protein homology).  
 GN NOV.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_Taxid=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129/SV; and ICR: TISSUE=Brain;  
 RX MEDLINE=97131523; PubMed=8975721;  
 RA Sneath M.R., Natarajan D., Taylor L.B., Choi C.P., Martinerie C.,  
 RA Perbal B., Schofield P.N., Boulter C.A.;  
 RT Genomic structure and chromosomal mapping of the mouse nov gene.";  
 RL Genomics 38:425-428(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6;  
 RX MEDLINE=96204003; PubMed=8622864;  
 RA Martinerie C., Chevalier G., Rauscher F.J. III, Perbal B.;  
 RT Regulation of nov by Wt1: a potential role for nov in  
 RT nephrogenesis.";  
 RL Oncogene 12:1479-1492(1996).  
 CC - FUNCTION: IMMEDIATE-EARLY PROTEIN LIKELY TO PLAY A ROLE IN CELL  
 CC GROWTH REGULATION (By Similarity).  
 CC - SUBUNIT: Interacts with FBLN1 (By similarity).  
 CC - SUBCELLULAR LOCATION: Secreted.  
 CC - SIMILARITY: Contains 1 IGFBP domain.  
 CC - SIMILARITY: Contains 1 VWF domain.  
 CC - SIMILARITY: Contains 1 TSP type-1 domain.  
 CC - SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>  
 CC or send an email to [license@isb-sdb.ch](mailto:license@isb-sdb.ch)).  
 CC  
 DR EMBL: X97863; CAA6457.1; -  
 DR EMBL: Y09257; CAA70454.1; -  
 DR EMBL: X96585; CAA65404.1; -  
 DR MGI: 109185; NOV.  
 DR InterPro: IPR006208; Cys\_knot.  
 DR InterPro: IPR006207; Cys\_knot\_C.  
 DR InterPro: IPR000867; Insl\_gro\_fac-pr.  
 DR InterPro: IPR000884; TSP1.  
 DR InterPro: IPR001007; VWF\_C.  
 DR Pfam: PF00007; Cys\_knot; 1.  
 DR Pfam: PF00219; IGFBP; 1.  
 DR Pfam: PF00090; TSP\_1; 1.  
 DR Pfam: PF00093; VWC; 1.  
 DR SMART: SM00041; CT; 1.  
 DR SMART: SM00121; IB; 1.  
 DR SMART: SM00209; TSP1; 1.  
 DR SMART: SM0214; VWC; 1.  
 DR PROSITE: PS01185; CTCK\_1; 1.  
 DR PROSITE: PS01225; CTCK\_2; 1.  
 DR PROSITE: PS00222; IGF\_BINDING; 1.  
 DR PROSITE: PS00922; TSP1; 1.  
 DR PROSITE: PS01208; VWF\_C\_1; 1.  
 DR PROSITE: PS0184; VWF\_C\_2; 1.  
 KW Growth factor; Signal.  
 FT SIGNAL 1 21  
 FT CHAIN 1 354  
 FT DOMAIN 29 101  
 FT DOMAIN 102 168  
 FT DOMAIN 202 247  
 FT DOMAIN 261 335  
 FT DISULFID 261 298  
 FT DISULFID 278 312  
 FT DISULFID 289 328  
 FT DISULFID 292 330  
 FT SEQUENCE 354 AA; 38667 MW; 717D9F85338028B96 CRC64; .)  
 Query Match 37.0%; Score 532.5; DB 1; Length 353;  
 Best Local Similarity 45.7%; Pred. No. 1.6e-33;  
 Matches 101; Conservative 32; Mismatches 73; Indels 15; Gaps 5;  
 26 CRTPC--TCPTPTPCPCPGVPLVLDGCGCCVCARLGGSCDHLVHCDPSGGLVQCPAG 83  
 33 CRRPCGGRCAPAPRCAPVPAVLDDGCGCLVCAQRGSCSPLLPCDSDSGGLYCDRGE 92  
 84 PGHGAVCLLDLDDSGCEVNGRRYLDGETFKPCRYLCRCDGFTCLPCLCSEDEVLPFM 143  
 93 DGGGIGICVLEGD--NCVFDMIRNGEFPQSCYKQCGRCRGQICLPRLNGLILPFP 151  
 144 DCPFRKRIQVPGKCCPEWVCDQ-----GVPAIORSAGHQLSALVTPASADAPCPN 196  
 152 DCPFRKRIQVPGKCCPEWVCDQ-----GVPAIORSAGHQLSALVTPASADAPCPN 196  
 153 DCPFRKRIQVPGKCCPEWVCDQ-----GVPAIORSAGHQLSALVTPASADAPCPN 196  
 197 WSTANGPCSTGGLGATRVSNONRFOLEIQRILCPPC 237  
 207 QTTEMSACSRSCGMGSHVTRVNRNOCCPEWVKTQRLCKMPPC 247

FT DISULFID 297 334 BY SIMILARITY.  
 FT CARBOHYD 91 91 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 277 277 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 354 AA; 38928 MW; 08CEBCEFC67829DE CRC64;  
 Query Match 36.8%; Score 529.5; DB 1; Length 354;  
 Best Local Similarity 42.5%; Pred. No. 2.7e-33;  
 Matches 105; Conservative 26; Mismatches 87; Indels 29; Gaps 6;  
 QY 10 LATSFLC--LISWCAQL---CRTPCPCMPFPPGCPGVPLVLDGCGCKXCARLGGSC 64  
 DB 10 LCLGFLFHLISQVSASLRCPSCPCPSISPCARVRSVLDGSCCPACARQEGSC 69  
 QY 65 DNLHVCPSQICQPCGAGPGHGAVALDDEDGSCENGRYLDGETFKPCNYLCRD 124  
 DB 70 SEMRCPQSSGLYCDRSDADPNNGICMVRPGD-NCVFDGYIRNGEKEFEPCYFCRCR 128  
 QY 125 DGGFTCLPLCESEYRLPSWDCPRPKRIQVPGKCCPEWCDGVPALQRTAGHQ--LS 182  
 DB 129 DGQIGCLPRCQLDVLLPGPCDPAIRKVAVPECECEKMTGC-----SDEGTGTG 179  
 QY 183 ALVTPA-----SADAPCPNMSTANGPCSTTCGLGIATRVSNQNFCELEIQR 230  
 DB 180 GLALPAIRPEATFVGEVSDSSINCIEQTEWMSACSKSCGMGVSTRTNRNQCEMXTQ 239  
 QY 231 LCLPRPC 237  
 DB 240 LCLVPRC 246  
 RESULT 13  
 NOV\_HUMAN STANDARD; PRT; 357 AA.  
 ID NOV\_HUMAN  
 AC PA8745: 096BVS;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE NOV protein homolog precursor (NOVH) (Nephroblastoma overexpressed  
 DE gene protein homolog).  
 GN NOV.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 OC NCBI\_TaxID:9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Placenta;  
 RX MEDLINE=94336229; PubMed=7520150;  
 RA Martinerie C., Huff V., Joubert I., Badziach M., Saunders G.F.,  
 RA Strong L.C., Perbal B.;  
 RT "Structural analysis of the human nov proto-oncogene and expression  
 RT in Wilms tumor";  
 RL Oncogene 9:2729-2732(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Jiang D., Gou D., Li W.;  
 RT "Cloning, sequencing and expression of human nov gene";  
 RL Submitted (MAR-2002) to the EMBL/Genbank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Skin;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunataratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahy J., Helton E., Kelleman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,  
 RA Scherf A., Schein J.E., Jones S.J.M., Maira M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [4]  
 RP INTERACTION WITH FBLN1.  
 RX PubMed=9927660;  
 RA Perbal B., Martinerie C., Salanson R., Werner M., He B., Rolzman B.;  
 RT "The C-terminal domain of the regulatory protein NOVH is sufficient to  
 RT promote interaction with fibulin 1C: a clue for a role of NOVH in  
 RT cell-adhesion signaling";  
 RL Proc. Natl. Acad. Sci. U.S.A. 96:869-874(1999).  
 CC -1- FUNCTION: IMMEDIATE-EARLY PROTEIN LIKELY TO PLAY A ROLE IN CELL  
 CC GROWTH REGULATION (BY SIMILARITY).  
 CC -1- SUBUNIT: Interacts with FBLN1.  
 CC -1- TISSUE SPECIFICITY: INCREASED EXPRESSION IN WILMS TUMOR OF THE  
 CC STROMAL TYPE.  
 CC -1- SIMILARITY: Contains 1 IGFBP domain.  
 CC -1- SIMILARITY: Contains 1 WFPC domain.  
 CC -1- SIMILARITY: Contains 1 TSP type-1 domain.  
 CC -1- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL: X78351; CAA55146.1; -;  
 DR EMBL: X78352; CAA55146.1; JOINED.  
 DR EMBL: X78353; CAA55146.1; JOINED.  
 DR EMBL: X78354; CAA55146.1; JOINED.  
 DR EMBL: X96584; CAA63403.1; -;  
 DR EMBL: AY082381; AAL92490.1; -;  
 DR EMBL: BC015028; AAH15028.1; -;  
 DR PIR: I38069; I38069.  
 DR GeneW: HGNC:7885; NOV.  
 DR MIM: 164958; -;  
 DR InterPro: IPR006208; Cys\_knot.  
 DR InterPro: IPR006207; Cys\_knot\_C.  
 DR InterPro: IPR000867; Ins1\_gro\_fac-pr.  
 DR InterPro: IPR000884; TSP1.  
 DR InterPro: IPR001007; WVF\_C.  
 DR Pfam: PF00007; Cys\_knot; 1.  
 DR Pfam: PF00219; IGFBP; 1.  
 DR Pfam: PF00090; TSP\_1; 1.  
 DR Pfam: PF00093; WVC; 1.  
 DR Pfam: SM00041; CT; 1.  
 DR SMART: SM00121; IB; 1.  
 DR SMART: SM00209; TSP1; 1.  
 DR SMART: SM00214; WVC; 1.  
 DR PROSITE: PS01185; CTCK\_1; 1.  
 DR PROSITE: PS01225; CTCK\_2; 1.  
 DR PROSITE: PS00222; IGF\_BINDING; 1.  
 DR PROSITE: PS00992; TSP1; 1.  
 DR PROSITE: PS01208; WVF\_C; 1.  
 DR PROSITE: PS0184; WVF\_C2; 1.  
 DR Growth factor; Signal.  
 KW SIGNAL  
 FT CHAIN 1 27  
 FT DOMAIN 28 357 POTENTIAL.  
 FT DOMAIN 35 107 NOV PROTEIN HOMOLOG.  
 FT DOMAIN 108 174 IGFBP.  
 FT DOMAIN 205 250 WVF\_C.  
 FT DOMAIN 264 338 TSP TYPE-1.  
 FT DISULFID 264 301 CTCK.  
 FT BY SIMILARITY.









GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 25, 2003, 12:32:15 ; Search time 96 Seconds  
(without alignments)  
672.012 Million cell updates/sec

Title: US-10-010-408-2  
Perfect score: 1440  
Sequence: 1 MRGSPHLHLATSFLLCLLSM.....LCPPPCPLARSHSWSNCAF 250

Scoring table:  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_23:\*  
1: sp-archaea:\*  
2: sp-bacteria:\*  
3: sp-fungi:\*  
4: sp-human:\*  
5: sp-invertebrate:\*  
6: sp-mammal:\*  
7: sp-mhc:\*  
8: sp-organelle:\*  
9: sp-phage:\*  
10: sp-plant:\*  
11: sp-rodent:\*  
12: sp-virus:\*  
13: sp-vertebrate:\*  
14: sp-unclassified:\*  
15: sp-virus:\*  
16: sp-bacteriap:\*  
17: sp-archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID | Description         |
|------------|--------|-------------|--------|-------|---------------------|
| 1          | 1302.5 | 90.5        | 251    | 11    | 08CIC8 mus musculu  |
| 2          | 575.5  | 40.0        | 347    | 13    | Q9PT80 notophthalm  |
| 3          | 551    | 38.3        | 344    | 13    | Q98T08 gallus galli |
| 4          | 550.5  | 38.2        | 343    | 13    | 042607 xenopus lae  |
| 5          | 529.5  | 36.8        | 354    | 11    | 08CA67 mus musculu  |
| 6          | 517.5  | 35.9        | 375    | 13    | 098T55 xenopus lae  |
| 7          | 505    | 35.1        | 349    | 6     | 097765 sus scrofa   |
| 8          | 503.5  | 33.0        | 367    | 11    | Q99P00 ratius norv  |
| 9          | 503.5  | 33.0        | 367    | 11    | 054775 mus musculu  |
| 10         | 503    | 34.9        | 379    | 11    | Q9WTM9 ratius norv  |
| 11         | 498.5  | 34.6        | 367    | 4     | 095388 homo sapien  |
| 12         | 455    | 31.6        | 176    | 13    | 09PS86 gallus galli |
| 13         | 372    | 25.8        | 331    | 4     | 095358 homo sapien  |
| 14         | 371.5  | 25.8        | 354    | 4     | 095389 homo sapien  |
| 15         | 355.5  | 24.7        | 334    | 4     | Q9UID7 homo sapien  |
| 16         | 303    | 21.0        | 128    | 11    | Q9R2C0 ratius norv  |

|    |       |      |      |    |        |                     |
|----|-------|------|------|----|--------|---------------------|
| 17 | 294.5 | 20.5 | 100  | 4  | Q9UDL6 | Q9UDL6 homo sapien  |
| 18 | 291   | 20.2 | 280  | 4  | Q9HCS3 | Q9HCS3 homo sapien  |
| 19 | 257   | 17.8 | 230  | 4  | Q8WYK7 | Q8WYK7 homo sapien  |
| 20 | 210   | 14.6 | 374  | 5  | Q9VVK0 | Q9VVK0 drosophila   |
| 21 | 210   | 14.6 | 470  | 5  | Q9VVK3 | Q9VVK3 drosophila   |
| 22 | 200   | 13.9 | 77   | 4  | Q9UDE4 | Q9UDE4 homo sapien  |
| 23 | 161.5 | 11.2 | 1664 | 5  | Q9VYQ2 | Q9VYQ2 caenorhabdi  |
| 24 | 161   | 11.2 | 626  | 4  | Q8ND91 | Q8ND91 homo sapien  |
| 25 | 161   | 11.2 | 969  | 4  | Q96KG6 | Q96KG6 homo sapien  |
| 26 | 158.5 | 11.0 | 1637 | 6  | Q9XSV8 | Q9XSV8 bos taurus   |
| 27 | 158.5 | 11.0 | 5146 | 6  | Q8SPM4 | Q8SPM4 bos taurus   |
| 28 | 156.5 | 10.9 | 4123 | 4  | Q75851 | Q75851 homo sapien  |
| 29 | 154.5 | 10.7 | 747  | 11 | Q8VHF4 | Q8VHF4 mus musculu  |
| 30 | 154.5 | 10.7 | 1034 | 11 | Q8VHF7 | Q8VHF7 mus musculu  |
| 31 | 153.5 | 10.7 | 1004 | 11 | Q8CGA7 | Q8CGA7 mus musculu  |
| 32 | 153.5 | 10.7 | 1034 | 11 | Q8VTK5 | Q8VTK5 mus musculu  |
| 33 | 150.5 | 10.5 | 1048 | 13 | Q8AWW5 | Q8AWW5 gallus galli |
| 34 | 149.5 | 10.4 | 2327 | 13 | Q9IBG7 | Q9IBG7 xenopus lae  |
| 35 | 149   | 10.3 | 1028 | 11 | Q9JLLO | Q9JLLO mus musculu  |
| 36 | 148.5 | 10.3 | 947  | 11 | Q8BKK7 | Q8BKK7 mus musculu  |
| 37 | 144   | 10.0 | 685  | 6  | Q9TTS5 | Q9TTS5 bos taurus   |
| 38 | 144   | 10.0 | 1246 | 4  | Q75095 | Q75095 homo sapien  |
| 39 | 139.5 | 9.7  | 406  | 11 | Q92513 | Q92513 mus musculu  |
| 40 | 139.5 | 9.7  | 426  | 11 | Q8VEA6 | Q8VEA6 mus musculu  |
| 41 | 139.5 | 9.7  | 482  | 11 | Q9CVG8 | Q9CVG8 mus musculu  |
| 42 | 139.5 | 9.7  | 483  | 11 | Q8K001 | Q8K001 mus musculu  |
| 43 | 139.5 | 9.7  | 938  | 11 | Q8CIA2 | Q8CIA2 mus musculu  |
| 44 | 139   | 9.7  | 425  | 6  | Q02661 | Q02661 bos taurus   |
| 45 | 139   | 9.7  | 507  | 5  | Q8WS81 | Q8WS81 drosophila   |

## ALIGNMENTS

| RESULT 1 | ID  | Q8CIC8                               | PRELIMINARY; | PRT; | 251 AA. |
|----------|---|--------------------------------------|--------------|------|---------|
| AC       | Q8CIC8  | 01-MAR-2003 (TREMBLrel. 23, Created) |              |      |         |
| DT       | 01-MAR-2003 (TREMBLrel. 23, Last sequence update)         |                                      |              |      |         |
| DT       | 01-MAR-2003 (TREMBLrel. 23, Last annotation update)       |                                      |              |      |         |
| DE       | Similar to WNT1 inducible signaling pathway protein 2.    |                                      |              |      |         |
| OS       | Mus musculus (Mouse).                                     |                                      |              |      |         |
| OC       | Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;   |                                      |              |      |         |
| OC       | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus. |                                      |              |      |         |
| OX       | NCBI_Taxid=10090;   |                                      |              |      |         |
| RN       | [1]   |                                      |              |      |         |
| RP       | SEQUENCE FROM N.A.  |                                      |              |      |         |
| RC       | STRAIN=C57BL/6J; TISSUE=Mammary gland;                    |                                      |              |      |         |
| RA       | Strusberg, R.;  |                                      |              |      |         |
| RL       | Submitted (JUN-2002) to the EMBL/Genbank/DBJ databases.   |                                      |              |      |         |
| DR       | EMBL; BC032877; AAH32877.1; --                            |                                      |              |      |         |
| SQ       | SEQUENCE 251 AA; 27106 MW; 57520309A9069524 CRC64;        |                                      |              |      |         |

| Query Match | Best Local Similarity | Score   | 90.5%; | DB 11;     | Length 251; |
|-------------|-----------------------|---|--------|------------|-------------|
| Matches     | 225;                  | Conservative  | 9;     | Mismatches | 16;         |
| Indels      | 1;                    | Gaps  | 1;     |            |             |
| QY          | 1                     | MRGSPHLHLATSFLLCLLSMWCAGQLCRPPCTCPMTPOCPQGVPLVLDGCGCCVCAARL 60  |        |            |             |
| DB          | 1                     | MRGNPLHLHLATSFLLCLLSMWCAGQLCRPPCTCPMTPOCPQGVPLVLDGCGCCVCAARL 60 |        |            |             |
| QY          | 61                    | GESCHLHVCDPSQGLVQPGAGPGHGAACVLLDDDDSGCEVNGRRYLDGTFKRCRYL 120    |        |            |             |
| DB          | 61                    | GESCHLHVCDPSQGLVQPGAGPGHGAACVLLDDDDSGCEVNGRRYLDGTFKRCRYL 120    |        |            |             |
| QY          | 121                   | CCDDGGFTCLPLCEDVRLPSMDCPRPRRIQVPCGCCPEWVCDDGV--TPAIRSTAGH 179   |        |            |             |
| DB          | 121                   | CCDDGGFTCLPLCEDVRLPSMDCPRPRRIQVPCGCCPEWVCDDGV--TPAIRSTAGH 179   |        |            |             |
| QY          | 180                   | QLSALVTPASADAPCPNMSTAWGPCSTTCGLGIATRVSNQNFQCLEIQRRLCLRPCLA 239  |        |            |             |
| DB          | 180                   | QLSALVTPASADAPCPNMSTAWGPCSTTCGLGIATRVSNQNFQCLEIQRRLCLRPCLA 239  |        |            |             |
| QY          | 181                   | QLSALVTPASADAPCPNMSTAWGPCSTTCGLGIATRVSNQNFQCLEIQRRLCLRPCLA 240  |        |            |             |
| DB          | 181                   | QLSALVTPASADAPCPNMSTAWGPCSTTCGLGIATRVSNQNFQCLEIQRRLCLRPCLA 240  |        |            |             |

```
OY 240 ARSHSSWNSAF 250
DB 241 SRSHGSMNSAF 251

RESULT 2
ID 09PT80 PRELIMINARY; PRT: 347 AA.
AC 09PT80;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Connective tissue growth factor.
GN CTGF.
OS Motophthalmus viridescens (Eastern newt) (Triturus viridescens).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Caudata; Salamandroidae; Salamandridae;
OC Motophthalmus.
OC NCBI_TaxID=8316;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Forelimb blastema;
RA Gates P.B.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Forelimb blastema;
RA Cash D.E., Gates P.B., Imokawa Y., Brookes J.P.;
RT "Identification of new connective tissue growth factor as a target of
RT retinoid regulation in limb blastemal cells."
RL Gene 222:119-124(1998).
DR EMBL: AJ271167; CAB65965.1;
DR Interpro: IPR006208; Cys_knot.
DR Interpro: IPR006207; Cys_knot_C.
DR Interpro: IPR000867; Insl_gro_fac-pr.
DR Interpro: IPR000884; TSPI.
DR Pfam: PF00007; Cys_knot.
DR Pfam: PF00219; IGFBP.
DR Pfam: PF00090; tsf_1;
DR Pfam: PF00093; vwc_1.
DR SMART: SM00041; CT; 1.
DR SMART: SM00121; IB; 1.
DR SMART: SM00209; TSPI; 1.
DR SMART: SM00214; VWC; 1.
DR PROSITE: PS01185; CTCK_1; 1.
DR PROSITE: PS01225; CTCK_2; 1.
DR PROSITE: PS00222; IGF_BINDING; 1.
DR PROSITE: PS50092; TSPI; 1.
DR PROSITE: PS01208; VWC; 1.
SQ SEQUENCE 347 AA; 38098 MW; 3B7E2399F27672C1 CRC64;

Query Match 40.0%; Score 575.5; DB 13; Length 347;
Best Local Similarity 47.5%; Pred. No. 2.6e-47;
Matches 113; Conservative 26; Mismatches 84; Indels 15; Gaps 4;

OY 9 LIAATFLCLLSWCAQLCTPCTPMTPPQCGQGVLYLVDDGGCCVCARRLGESDHLH 68
DB 13 LIAVALLSWVS--CAQDGSCECRCPNKPPECAGISLVWDGCGCCVCAKQJGEICTED 70
OY 69 VCDPSQGLVCPGAGPGGHGAVCLLDEDDGSCENVGRYRNDGETFKPNCRVLCRDDGGF 128
DB 71 VCDPFRGLFCDFGSRNKKIGVCTA-KDGAFCVFGMYRSGESFSSCKYQCTCLDGGV 129
OY 129 TCLPLCSEEDVRLPSMDCPPRKRIQVPGKCCPEWVDO-----GVTPAIQSTAGHOL 181
DB 130 GCVPLCGVDVRLPSDCEPPRRVKQKGCCEWVDOPEQRTRVSPALAVYRQEEETGPD 189
OY 182 SALVTPASADAPCPMNSTAWGSCSTTCGIGIATRVSNONRFFCOLIQRILCPRLA 239
DB 190 SSLMR-----ANCLVOTTEMSACSTCKGIGISTRTVNDNMCRLKQSRILCAVRCEA 242
```

```
RESULT 3
ID 098T08 PRELIMINARY; PRT: 344 AA.
AC 098T08;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Connective tissue growth factor precursor (connective tissue growth
DE factor/hypertrophic chondrocyte-specific protein 24).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OC NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Gygi D.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Mukudai Y., Kubota S., Takigawa M.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ298335; CAC33438.1;
DR EMBL: AF463517; AAL68834.1;
DR Interpro: IPR006208; Cys_knot.
DR Interpro: IPR006207; Cys_knot_C.
DR Interpro: IPR000867; Insl_gro_fac-pr.
DR Interpro: IPR000884; TSPI.
DR Interpro: IPR001007; VWF_C.
DR Pfam: PF00007; Cys_knot; 1.
DR Pfam: PF00219; IGFBP; 1.
DR Pfam: PF00090; tsf_1; 1.
DR Pfam: PF00093; vwc_1.
DR SMART: SM00041; CT; 1.
DR SMART: SM00121; IB; 1.
DR SMART: SM00209; TSPI; 1.
DR SMART: SM00214; VWC; 1.
DR PROSITE: PS01185; CTCK_1; 1.
DR PROSITE: PS01225; CTCK_2; 1.
DR PROSITE: PS00222; IGF_BINDING; 1.
DR PROSITE: PS50092; TSPI; 1.
DR PROSITE: PS01208; VWC; 1.
DR SIGNAL.
FT CHAIN 1 21 POTENTIAL.
FT SIGNAL 22 344 CONNECTIVE TISSUE GROWTH FACTOR.
SQ SEQUENCE 344 AA; 37499 MW; 69E639AF65BF1D00 CRC64;

Query Match 38.3%; Score 551; DB 13; Length 344;
Best Local Similarity 45.5%; Pred. No. 5.8e-45;
Matches 111; Conservative 25; Mismatches 94; Indels 14; Gaps 4;

OY 4 SPLIHLATSLFLCLLSWCAQLCTPCTPMTPPQCGQGVLYLVDDGGCCVCARRLGE 62
DB 2 SPASIALVALLLALLPKRGQSGCQCCGSGPGSPGVSILVDDGGCCVCARRLGE 61
OY 63 SCDHLVCDPSQGLVCPGAGPGGHGAVCLLDEDDGSCENVGRYRNDGETFKPNCRVLCR 122
DB 62 LCTERDPCDHHGGLCDGDSPPANRRIQVCTA-RDGAFCVFSAMYRSGESFSSCKYQCT 120
OY 123 CDDGFTCLPLCSEEDVRLPSMDCPPRKRIQVPGKCCPEWVDOGVTPAIQSTAGHOL 182
DB 121 CUDGAVGCVPLCSMDVRLPSDCEPPRRVKLGKCEWVDEA-----KEQTAVGPALA 175
OY 183 AL-----VTPASADAPCPMNSTAWGSCSTTCGIGIATRVSNONRFFCOLIQRILCP 235
DB 176 ARLEEDTYGPDPTMNRANCLVQTTETMSACSTCKGIGISTRTVNDNMCRLKQSRILCAVR 235
OY 236 PCLA 239
DB 236 PCEA 239
```

```

RESULT 4
ID 042607 PRELIMINARY: PRT: 343 AA.
AC 042607
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Connective tissue growth factor XCTGF.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Xing Z., King M.L.;
RL Submitted (Aug-1997) to the EMBL/Genbank/DBJ databases.
DR EMBL: U43524; AAB67639.1; -
DR EMBL: U43523; AAB67638.1; -
DR InterPro: IPR006208; Cys_knot.C.
DR InterPro: IPR006207; Cys_knot.C.
DR InterPro: IPR00867; Insl_gro_fac-pr.
DR InterPro: IPR00884; TSP1.
DR InterPro: IPR001007; VWF_C.
DR Pfam: PF00007; Cys_knot; 1.
DR Pfam: PF00219; IGFBP; 1.
DR Pfam: PF00090; tsp_1; 1.
DR Pfam: PF00093; vwc; 1.
DR SMART: SM00041; CT; 1.
DR SMART: SM00121; IB; 1.
DR SMART: SM00209; TSP1; 1.
DR SMART: SM00214; VMC; 1.
DR PROSITE: PS01185; CTCK_1; 1.
DR PROSITE: PS01225; CTCK_2; 1.
DR PROSITE: PS50092; TSP1; 1.
DR PROSITE: PS01208; VWF; 1.
SQ SEQUENCE 343 AA; 37966 MW; 93F221C5DB565A81 CRC64;

Query Match 38.2%; Score 550.5; DB 13; Length 343;
Best Local Similarity 45.7%; Pred. No. 6; 5e-45;
Matches 107; Conservative 26; Mismatches 84; Indels 17; Gaps 4;

OY 17 LLSWVC---AQLCTPCTCPMTPPQCPGVPVLVDGCGCCVKCARLGESCDHLHVCDP 72
DB 11 LFAICWVSDAOECGECQCPMKYVCDPGVAMVDGCGCCVKSGELGELCTERDVCDP 70
OY 73 SGLVCPQAPGPGHGAVALDDEDDGSCENVNGRRYLDGETFKPNCRVLCRCDDGGFTCLP 132
DB 71 HKGLFCDFGSRYNRKIGVCTARE-GAPCVFGSTVYRSGESFQSSCKYCTCIDGSGVCVP 129
OY 133 LCSEVRLPSMDCPPRKRIQVPGKCCPEWVCDQGYTAIONSTAOGHLSAL-----V 185
DB 130 LCSMDIRIPSPCEPPRRVKLPKCCCEWVCDQ-----QERTLGPALPFRMETYGP 184
OY 186 TPASADACPMWSTAMGPGSTTCGLGIATRVSNONRFOLEIQRRLCLPRCLA 239
DB 185 DPSLRANCLVOTTEWSKSTCGMGISTRTVINDNEHCKLEKQSRKLCVAPCEA 238

RESULT 5
ID 08CA67 PRELIMINARY: PRT: 354 AA.
AC 08CA67
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Nephroblastoma overexpressed.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; Tissue=Spinal cord;
RX MEDLINE=22354683; Pubmed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL 60,770 full-length cDNAs."
DR EMBL: AK039481; BAC30363.1; -
DR EMBL: AK039481; BAC30363.1; -
SQ SEQUENCE 354 AA; 38828 MW; 6F3B34A3A84187EE CRC64;

Query Match 36.8%; Score 529.5; DB 11; Length 354;
Best Local Similarity 42.5%; Pred. No. 7e-43;
Matches 105; Conservative 26; Mismatches 87; Indels 29; Gaps 6;

OY 10 LATSFLC--LISWCAQL---CRPPCTCPMTPPQCPGVPVLVDGCGCCVKCARLGESC 64
DB 10 LCLGFLFHLISQVSAISLRCPSPISPTCAPGVRVLDGSCCPVCAARQREGSC 69
OY 65 DHLHVCDSGIVCPGAPGPGHGAVALDDEDDGSCENVNGRRYLDGETFKPNCRVLCRCD 124
DB 70 SEMRCDSSGLYCDRSADPNNGTICWVPEGD-NCVFDGYIYRNGEYFERNQYFCTCR 128
OY 125 DGGFTCLPLCSEEDYRLPSMDCPPRKRIQVPGKCCPEWVCDQGYTAIONSTAOGHQ--LS 182
DB 129 DGIQGLCPRLQCLDVLTPGDCPAPRKVAVPGCCCKMTG-----SDEGNGQGLG 179
OY 183 ALVTPA-----SADAPCPMSTAMGPGSTTCGLGIATRVSNONRFOLEIQR 230
DB 180 GLALPAPPEATVGEVSDSSINCIQETEMWSACSKSGMGVSTFVTRNRNCKEAVKOTR 239
OY 231 LCLPRPC 237
DB 240 LCLVAPC 246

RESULT 6
ID 098TX5 PRELIMINARY: PRT: 375 AA.
AC 098TX5
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE secreted cysteine-rich protein cyr61.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Latinkic B.V., Bennett B., Smith J.C.;
RT "Characterization of Xenopus cyr61."
RL Submitted (NOV-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF320592; AAK0947.1; -
DR InterPro: IPR006208; Cys_knot.
DR InterPro: IPR006207; Cys_knot.C.
DR InterPro: IPR00867; Insl_gro_fac-pr.
DR InterPro: IPR00884; TSP1.
DR InterPro: IPR001007; VWF_C.
DR Pfam: PF00007; Cys_knot; 1.
DR Pfam: PF00219; IGFBP; 1.
DR Pfam: PF00090; tsp_1; 1.
DR Pfam: PF00093; vwc; 1.
DR SMART: SM00041; CT; 1.
DR SMART: SM00121; IB; 1.
DR SMART: SM00209; TSP1; 1.
DR SMART: SM00214; VMC; 1.
DR PROSITE: PS01185; CTCK_1; 1.
DR PROSITE: PS01225; CTCK_2; 1.
DR PROSITE: PS00222; IGF_BINDING; 1.
DR PROSITE: PS50092; TSP1; 1.
DR PROSITE: PS01208; VWF; 1.

```

```

SQ SEQUENCE 375 AA; 41460 MW; 78075CA7B380304E CRC64;
Query Match 35.9%; Score 517.5; DB 13; Length 375;
Best Local Similarity 37.2%; Pred. No. 1.1e-41;
Matches 99; Conservative 39; Mismatches 87; Indels 41; Gaps 5;

QY 13 SFLLCLSMVCAQL-----CRPTCTCPWTPPOGVPVLVDGCGCKVCARRLGG 61
DB 2 SFLLANPVLALALSGFIDLAIVSSCPAVCCPEVPCAPAGVLVDGCGCKKICAKQLN 61
QY 62 ESCDHLVHVDPSGGLGCPGAGPGGAGVAVCLDEDDGSCGVNGRRYLDGETRPNCRVLC 121
DB 62 EDCSKHPDHTKGLKCNFGASRAIKGICRAKSEGRPEVNSKITQNGESFOPNCKHOC 121
QY 122 RCDGGLFTCLPLCSEVDRLPSWDCPRKRIQVPGKCEPMVCDQGVTPAIQ----- 172
DB 122 TCIDGAVGCLPLCPGELSLPNLGCPRVLRVVGCCCEMVCDKADPDVDEMDFFNKEF 181
QY 173 -RSTAOG-----HQLSALV-----TPAS--ADAPCPMNSTAMGPGSTTCGLG 211
DB 182 GMDTNEGELTRKNEFAVVKGLKMLPVGSDPDSHVNSKCIQVOTSMSCSKTCGTG 241
QY 212 IATRVSNORFCOLEIQRRLCLPRPC 237
DB 242 ISTRVTNDNSNCRVLAETRICEVRC 267

RESULT 7
O97765 PRELIMINARY; PRT; 349 AA.
AC 097765;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DE 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Connective tissue growth factor.
GN CTGF.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN 11
RP SEQUENCE FROM N.A.
RC TISSUE=uterus;
RA Harding P.A., Bristolstock D.R.;
RT "Cloning and sequencing of a porcine connective tissue growth factor
(CTGF) cDNA."
RL Submitted (SEP-1996) to the EMBL/Genbank/DBJ databases.
DR EMBL: U70060; AAD00174.1; -
DR InterPro: IPR006208; Cys_knot.
DR InterPro: IPR006207; Cys_knot_C.
DR InterPro: IPR000867; Insl_gro_fac_pr.
DR InterPro: IPR000884; TSPL.
DR InterPro: IPR001007; WMF_C.
DR Pfam: PF00007; Cys_knot. 1.
DR Pfam: PF000219; IGFBRP. 1.
DR Pfam: PF00090; TSP_1. 1.
DR Pfam: PF00093; vwc. 1.
DR Pfam: PF00093; vwc. 1.
DR SMART: SM00041; CT. 1.
DR SMART: SM00121; IB. 1.
DR SMART: SM00209; TSP1. 1.
DR SMART: SM00214; vwc. 1.
DR PROSITE: PS01185; CTCK_1. 1.
DR PROSITE: PS01225; CTCK_2. 1.
DR PROSITE: PS00222; IGF_BINDING. 1.
DR PROSITE: PS00922; TSP1_BINDING. 1.
DR PROSITE: PS01208; WMFC. 1.
SQ SEQUENCE 349 AA; 37946 MW; 35AB4275AC1DAB3A CRC64;

Query Match 35.1%; Score 505; DB 6; Length 349;
Best Local Similarity 42.3%; Pred. No. 1.6e-40;
Matches 101; Conservative 32; Mismatches 88; Indels 18; Gaps 5;

13 SFLLCLSM-----VCAQLCRPTCTCPWTPPO-CPGCVPLVDGCGCKVCARRLGGSCHDL 67

```

```

DB 12 AFVLLALICSRPASGDDCGGCOCAAGKRACPAVSSILEGCGCRLCAKHLGDLCTER 71
QY 68 HVCDDPSGGLVCPGAGPGGHAIVCLDEDDGSCGVNGRRYLDGETRPNCRVLCRCDG 127
DB 72 APCDDPKHGLPCPGSAPANKIAGCTA-KCAPCVGGVYKSGESQSSCKTQCTCLDGA 130
QY 128 FTCLPLCSEVDRLPSWDCPRKRIQVPGKCEPMVCDQGVTPAIQSTAGHQLSAL--- 184
DB 131 VCGVPLCSMDVRLPSDPCFPFRVRLPKKCEEMVCDDEP-----KDHVYGPALAAVRL 185
QY 185 -----VTPASADAPCPMNSTAMGPGSTTCGLGATRVSNORFCOLEIQRRLCLPR 239
DB 186 DTFGDPPTMNRANCLVOTTEWSAGSKTCMGISTRTVNDNASCRLEKOSRLCMVRCEA 244

RESULT 8
O99PPO PRELIMINARY; PRT; 367 AA.
AC 099PPO;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE ELM1.
GN ELM1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN 11
RP SEQUENCE FROM N.A.
RC MEDLINE=20487548; PubMed=11031104;
RA Slegman M.A., Murison J.G., Strachan L., Kumble K.D., Glenn M.P.,
RA McGrath A., Bickerstaff P., Grierson A., Havukkala I., Tan P.,
RA Watson J.D.;
RT "Gene expression in rat dermal papilla cells: analysis of 2529 ESTs."
RL Genomics 65:214-224(2000).
DR EMBL: AF28049; AAK00729.1; -
DR InterPro: IPR006208; Cys_knot.
DR InterPro: IPR006207; Cys_knot_C.
DR InterPro: IPR000867; Insl_gro_fac_pr.
DR InterPro: IPR000884; TSPL.
DR InterPro: IPR001007; WMF_C.
DR Pfam: PF00007; Cys_knot. 1.
DR Pfam: PF000219; IGFBRP. 1.
DR Pfam: PF00090; TSP_1. 1.
DR Pfam: PF00093; vwc. 1.
DR SMART: SM00041; CT. 1.
DR SMART: SM00121; IB. 1.
DR SMART: SM00209; TSP1. 1.
DR SMART: SM00214; vwc. 1.
DR PROSITE: PS01185; CTCK_1. 1.
DR PROSITE: PS01225; CTCK_2. 1.
DR PROSITE: PS00922; TSP1_BINDING. 1.
DR PROSITE: PS01208; WMFC. 1.
SQ SEQUENCE 367 AA; 40613 MW; 8AA434C69D3243D2 CRC64;

Query Match 35.0%; Score 503.5; DB 11; Length 367;
Best Local Similarity 43.7%; Pred. No. 2.3e-40;
Matches 94; Conservative 29; Mismatches 89; Indels 3; Gaps 3;

QY 24 QLCRPTCTCPWTPPOGVPVLVDGCGCKVCARRLGGSCHDLHVCDDPSGGLVCPG 82
DB 47 EFCWKECECPQAPRPRCLVSLITDGCECKICAOOLIGNCTEAAVCDPRHLXCDYSGD 106
QY 83 GGGGAGVCLLDEDDGSCGVNGRRYLDGETRPNCRVLCRCDGGETCLPLCSEVDRLPS 142
DB 107 RPRVAGVCAQVYGVG-CVLDSVRYTNGSFQPNCRYNCTCIDGTGVCPPILSP-RPR 164
QY 143 WCCPRPRKRIQVPGKCEPMVCDQGVTPAIQSTAGHQLSALVTPASADAPCPMNSTAM 202
DB 165 LMCGRPRHVRVVGCCCEMVCDKADARPRNQTALDLTTRAFASGAVDQRENCIATVSPMS 224

```



OY 210 LGIATRVSNQNFQLEIQRILCLPRPC 237  
 DB 243 TGISTRTVNDNECRKLVETKRICVPRPC 270

## RESULT 11

OY95388 PRELIMINARY: PRT: 367 AA.  
 AC 095388;  
 DT 01-MAY-1999 (TREMBlrel. 10, Created)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)  
 DT 01-MAY-2003 (TREMBlrel. 23, Last annotation update)  
 DE 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
 GN WISP1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 OX NCBI\_Taxid=9606;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Lung, and Fetal Kidney;  
 RX MEDLINE=99061933; PubMed=9843955;  
 RA Penica D., Swanson T.A., Welsh J.W., Roy M.A., Lawrence D.A., Lee J.,  
 RA Brush J., Taneyhill L.A., Deuel B., Lew M., Watanabe C., Cohen R.L.,  
 RA Mellam A.F., Finley G.G., Quilke P., Goddard A.D., Hillan K.J.,  
 RA Gurley A.L., Botstein D., Levine A.J.;  
 RT "WISP genes are members of the connective tissue growth factor family  
 RT that are up-regulated in wnt-1-transformed cells (and aberrantly  
 RT expressed in human colon tumors)."  
 RL Proc. Natl. Acad. Sci. U.S.A. 95:14717-14722(1998).  
 RN (12)  
 RP SEQUENCE FROM N.A.  
 RA Blechschmidt K., Kalaydjieva L., Goodman R., Gresham D., Baas F.,  
 RA Jonge R.d., Schilhabel M., Schattevoy R., Dettle M., Menzel U.,  
 RA Rosenthal A.;  
 RL Submitted (Oct-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL, AF100779; AAC96321.1;  
 DR EMBL, AF192304; AAF2341.1;  
 DR Genew; HGNC:12769; WISP1.  
 DR InterPro: IPR006208; Cys\_knot.  
 DR InterPro: IPR006207; Cys\_knot.  
 DR InterPro: IPR00867; Insl\_gro\_fac-pr.  
 DR InterPro: IPR00884; TSP1.  
 DR InterPro: IPR001007; VWF-C.  
 DR Pfam: PF000219; IGFBP; 1.  
 DR Pfam: PF00090; TSP1; 1.  
 DR Pfam: PF00093; VWC; 1.  
 DR SMART; SM00041; CT; 1.  
 DR SMART; SM00121; IB; 1.  
 DR SMART; SM00209; TSP1; 1.  
 DR SMART; SM00214; VWC; 1.  
 DR PROSITE; PS01185; CTCK\_1; 1.  
 DR PROSITE; PS01225; CTCK\_2; 1.  
 DR PROSITE; PS0092; TSP1; 1.  
 DR PROSITE; PS01208; VWF-C; 1.  
 SQ SEQUENCE 367 AA; 40331 MW; 9F29CA94D69C0502 CRC64;

Query Match 34.6%; Score 498.5; DB 4; Length 367;  
 Best Local Similarity 43.8%; Pred. No. 7e-40; 81; Indels 11; Gaps 7;  
 Matches 96; Conservative 31; Mismatches 81;

OY 24 QLCRTPTCPMTPOGPGVPLVLDGCGCCCKVCARRLESCDHLHVCPSGLVQCGA 82  
 DB 47 QFCMKPCPCSPSPRCPLVSLITDGCCECKMKCAQGLDNDGNEALICDPHRLYLDYSGD 106  
 OY 83 GPGGAGVCLLDEDDGSCVENGRYLDGTEFKPNCVLCRCDGFTCLPLCSEDEVRLPS 142  
 DB 107 RPRVATGCAQVGVG-CVLGDVRYNNGOSFQPNCKYKNCOTGDGAVGCTPLCLR-VPRPR 164  
 OY 143 WDCPRKRIQVPGKCCPEWVC-DQGVTPAIQRTSTQGHQSLVTPASADA---PCPPWS 198

DB 165 LMCPRRVSIPGHCCSQWVCEDDAKRP---RKTAP-RDTGAFDAVGEVAMHNRNCIAYT 220  
 OY 199 TAMGPGSTTCGLIATRVSNQNFQLEIQRILCLPRPC 237  
 DB 221 SPMSPCSTSCGLGVSTRTISNVNAOCMPEDESRLCLRRC 259

## RESULT 12

OY95388 PRELIMINARY: PRT: 176 AA.  
 AC 095388;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)  
 DE Potential IGF binding protein (Fragments).  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_Taxid=9031;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92096871; PubMed=1756408;  
 RA Martinerie C., Perbal B.;  
 RT "Expression of a gene encoding a novel potential IGF binding protein  
 RT in human tissues."  
 RL C. R. Acad. Sci., III, Sci. Vie 313:345-351(1991).  
 DR InterPro: IPR00867; Insl\_gro\_fac-pr.  
 DR InterPro: IPR001007; VWF-C.  
 DR SMART; SM00121; IB; 1.  
 DR SMART; SM00214; VWC; 1.  
 DR PROSITE; PS00222; IGF\_BINDING; 1.  
 DR PROSITE; PS01208; VWF-C; 1.  
 FT NON\_TER 1 1  
 FT NON\_CONS 77 78  
 FT NON\_CONS 154 155  
 FT NON\_TER 176 176  
 SQ SEQUENCE 176 AA; 18656 MW; 4FA69FFEDB79B1C67 CRC64;

Query Match 31.6%; Score 455; DB 13; Length 176;  
 Best Local Similarity 42.1%; Pred. No. 5e-36; 59; Indels 32; Gaps 3;  
 Matches 83; Conservative 23; Mismatches 59;

OY 26 CRTPC--TCPMTPOGPGVPLVLDGCGCCCKVCARRLESCDHLHVCPSGLVQCGAG 83  
 DB 10 CPRPCGRCAPAPRCARCAVPALVLDGCGCCCLVCARQRESCSPLPLPCSSGGLVCDRQPE 69  
 OY 84 PGHGAVCLLDEDDGSCVENGRYLDGTEFKPNCVLCRCDGFTCLPLCSEDEVRLPSW 143  
 DB 70 DGGGAGIC---EGDNCFVDEGMVYRNGEFPQSCYQCTCRDQIGCLPRCNLGLLPGP 125  
 OY 144 DCPRRKRIQVPGKCCPEWVCDOGVTPAIQRTSTQGHQSLVTPASADAPCPNMSTWAGP 203  
 DB 126 DCPFRKRIEVEGCECKMKVCDPRDEVL-----GTENSA 159  
 OY 204 CSTTGIGIATRVSNQNF 220  
 DB 160 CSKSCGMGFSFRTNRN 176

## RESULT 13

OY95388 PRELIMINARY: PRT: 331 AA.

AC 095388;  
 DT 01-MAY-1999 (TREMBlrel. 10, Created)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)  
 DT 01-MAY-2003 (TREMBlrel. 23, Last annotation update)  
 DE DJ142L.3 (Connective tissue growth factor (NOV, GIG) like protein)  
 GN (Fragment).  
 GN DJ142L.3 OR LIIC.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.





DR SMART; SM00041; CT; 1.  
 DR SMART; SM00209; TSP1; 1.  
 DR SMART; SM00214; VWC; 1.  
 DR PROSITE; PS01225; CTCK\_2; 1.  
 DR PROSITE; PS50092; TSP1; 1.  
 DR PROSITE; PS01208; VWC; 1.  
 DR PROSITE; PS01208; VWC; 1.  
 SQ SEQUENCE 334 AA; 37246 MW; 9188987A7352E948; CRC64;

Query Match 24.7%; Score 355.5; DB 48 Length 334;  
 Best Local Similarity 30.3%; Pred. No. 3.6e-26;  
 Matches 74; Conservative 29; Mismatches 66; Indels 75; Gaps 5;

QY 49 GCGCCVCAARRLGECDL-HVCDP-----SGLVCOAGAPGCHGAVCLDEDDGSC 101  
 DB 2 GCGTHP-----NLCTHGHGTASPTSYKHHHTKGLECNAGASTALKGICRAOSEGRPCE 54  
 QY 102 VNGRRYLDGETFKPNCRVLCRDDDGFTCLPLCSEDRPSMDCCPRKRIQVPGKCCPEM 161  
 DB 55 YNSRIYQNGESFQPNCKHOCCTCIDGAVGCIPLCPQELSPNLCPPNRLVKYTGCCCEEM 114  
 QY 162 VCDQ-----GTPAIOR 173  
 DB 115 VDEDESIKPMEDQGLLGEKELGFDASEVELTRNNELIAGKSSLIKRIPVFGMEPRIRY 174  
 QY 174 STAQGHQLSALVTPASADAPCPNMSTAWGPCSTTCGLGATRVSNQNRFCOLEIORLCL 233  
 DB 175 NPLGQK-----CIQVQTSMSQCSKTCGTGISTRTNDNPECRLVKETPRICE 221  
 QY 234 PRPC 237  
 DB 222 VRPC 225

Search completed: July 25, 2003, 12:37:14  
 Job time : 98 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 15:59:01 ; Search time 2092.77 Seconds

(without alignments)  
104771.464 Million cell updates/sec

Title: US-10-010-408-3

Perfect score: 753  
Sequence: 1 ATGAGGGCAGCCACTGAT.....CATGACAGTGTCTTCTAA 753Scoring table: IDENTITY\_NUC  
Gapop 10.0' , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Database :

GenEmbl:\*

1: gb\_ba:\*

2: gb\_htg:\*

3: gb\_in:\*

4: gb\_com:\*

5: gb\_ov:\*

6: gb\_pat:\*

7: gb\_ph:\*

8: gb\_pl:\*

9: gb\_pr:\*

10: gb\_ro:\*

11: gb\_sts:\*

12: gb\_sy:\*

13: gb\_un:\*

14: gb\_vi:\*

15: em\_ba:\*

16: em\_fun:\*

17: em\_hum:\*

18: em\_in:\*

19: em\_mu:\*

20: em\_com:\*

21: em\_or:\*

22: em\_ov:\*

23: em\_pat:\*

24: em\_ph:\*

25: em\_pl:\*

26: em\_ro:\*

27: em\_sts:\*

28: em\_un:\*

29: em\_vi:\*

30: em\_htg\_hum:\*

31: em\_htg\_inv:\*

32: em\_htg\_other:\*

33: em\_htg\_mus:\*

34: em\_htg\_pln:\*

35: em\_htg\_rnd:\*

36: em\_htg\_mam:\*

37: em\_htg\_vrt:\*

38: em\_sy:\*

39: em\_htgo\_hum:\*

40: em\_htgo\_mus:\*

41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID         | Description         |
|------------|-------|-------------|--------|---------------|---------------------|
| 1          | 732.2 | 97.2        | 1741   | 10 AF259981   | AF259981 Rattus no  |
| 2          | 659   | 87.5        | 1734   | 6 AR210324    | AR210324 Sequence   |
| 3          | 659   | 87.5        | 1734   | 6 AR210325    | AR210325 Sequence   |
| 4          | 659   | 87.5        | 1734   | 10 AF100778   | AF100778 Mus muscu  |
| 5          | 654.2 | 86.9        | 1739   | 10 AF126063   | AF126063 Mus muscu  |
| 6          | 510.4 | 67.8        | 1266   | 6 AX076919    | AX076919 Sequence   |
| 7          | 510.4 | 67.8        | 1266   | 6 AX464186    | AX464186 Sequence   |
| 8          | 510.4 | 67.8        | 1283   | 6 AF083500    | AF083500 Homo sapi  |
| 9          | 510.4 | 67.8        | 1293   | 6 AR210322    | AR210322 Sequence   |
| 10         | 510.4 | 67.8        | 1293   | 6 AR210323    | AR210323 Sequence   |
| 11         | 510.4 | 67.8        | 1309   | 9 AF074604    | AF074604 Homo sapi  |
| 12         | 510.4 | 67.8        | 1427   | 9 AF100780    | AF100780 Homo sapi  |
| 13         | 510.4 | 67.8        | 1450   | 9 BC017782    | BC017782 Homo sapi  |
| 14         | 501.2 | 66.6        | 738    | 6 AR210337    | AR210337 Sequence   |
| 15         | 496.4 | 65.9        | 841    | 6 AR210338    | AR210338 Sequence   |
| 16         | 256.6 | 34.1        | 137964 | 2 AC126895    | AC126895 Rattus no  |
| 17         | 223.8 | 29.7        | 137964 | 2 AC126895    | AC126895 Rattus no  |
| 18         | 223.8 | 29.7        | 226303 | 2 AC095418    | AC095418 Rattus no  |
| 19         | 216.2 | 28.7        | 61072  | 10 AL731698   | AL731698 Mouse DNA  |
| 20         | 200.4 | 26.6        | 216757 | 2 AL669906    | AL669906 Mus muscu  |
| 21         | 169.8 | 22.5        | 107260 | 9 AL139352    | AL139352 Human DNA  |
| 22         | 168   | 22.3        | 1585   | 5 NV1271167   | NV1271167 Notophtha |
| 23         | 163   | 21.6        | 2075   | 6 AR018957    | AR018957 Sequence   |
| 24         | 163   | 21.6        | 2075   | 6 AR118595    | AR118595 Sequence   |
| 25         | 163   | 21.6        | 2075   | 6 AR130891    | AR130891 Sequence   |
| 26         | 163   | 21.6        | 2075   | 6 AX206708    | AX206708 Sequence   |
| 27         | 163   | 21.6        | 2075   | 6 I11636      | I11636 Sequence     |
| 28         | 163   | 21.6        | 2075   | 6 I32210      | I32210 Sequence     |
| 29         | 163   | 21.6        | 2075   | 9 HUMCONGRO   | HUMCONGRO           |
| 30         | 163   | 21.6        | 2312   | 9 HSCCTGP     | HSCCTGP             |
| 31         | 163   | 21.6        | 2998   | 6 AR119211    | AR119211 Sequence   |
| 32         | 163   | 21.6        | 2998   | 6 AR151276    | AR151276 Sequence   |
| 33         | 161   | 21.4        | 2267   | 6 AX206706    | AX206706 Sequence   |
| 34         | 161   | 21.4        | 2267   | 10 MUSFISPI2B | MUSFISPI2B          |
| 35         | 161   | 21.4        | 2330   | 10 MUSTGFB    | MUSTGFB             |
| 36         | 161   | 21.4        | 2334   | 10 BC006783   | BC006783 Mus muscu  |
| 37         | 159.2 | 21.1        | 1496   | 4 SSU83916    | SSU83916 Sus scrofa |
| 38         | 158   | 21.0        | 2338   | 6 E37595      | E37595 Monoclonal   |
| 39         | 158   | 21.0        | 2338   | 10 AB023068   | AB023068 Rattus no  |
| 40         | 158   | 21.0        | 2345   | 6 AX402019    | AX402019 Sequence   |
| 41         | 158   | 21.0        | 2345   | 10 AF120275   | AF120275 Rattus no  |
| 42         | 156.4 | 20.8        | 2350   | 6 AR194011    | AR194011 Sequence   |
| 43         | 156   | 20.7        | 2190   | 5 AR463517    | AR463517 Gallus ga  |
| 44         | 156   | 20.7        | 2288   | 5 GGA298335   | GGA298335 Gallus ga |
| 45         | 154.8 | 20.6        | 2350   | 6 AR201286    | AR201286 Sequence   |

## ALIGNMENTS

RESULT 1  
AF259981  
LOCUS  
DEFINITION Rattus norvegicus CCN family protein COP-1 (COP-1) mRNA, complete

ACCESSION AF259981  
VERSION AF259981.1  
KEYWORDS GI:7739780  
SOURCE  
ORGANISM Rattus norvegicus  
Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
REFERENCE 1 (bases 1 to 1741)  
AUTHORS Zhang, R., Averboukh, L., Zhu, W., Zhang, H., Jo, H., Dempsey, P. J.,

TITLE Coffey, R.J., Pardee, A.B. and Liang, P.  
Identification of rCop-1, a new member of the CCN protein family,  
as a negative regulator for cell transformation  
JOURNAL Mol. Cell. Biol. 18 (10), 6131-6141 (1998)  
MEDLINE 98414629  
PUBMED 9742130  
REFERENCE 2 (bases 1 to 1741)  
AUTHORS Liang, P.  
JOURNAL Direct Submission  
TITLE Submitted (24-APR-2000) Cell Biology, Vanderbilt-Ingram Cancer  
Center, 649 MRB II, Nashville, TN 37232, USA  
FEATURES  
source Location/Qualifiers  
1..1741  
/organism="Rattus norvegicus"  
/db\_xref="taxon:10116"  
1..1741  
/gene="Cop-1"  
262..1014  
/note="secreted protein"  
/product="CCN family protein COP-1"  
/codon\_start=1  
/protein\_id="AA69011.1"  
/db\_xref="GI:7739781"  
/translation="MRGSPILRLATSLFLCLSNVCAQLCTPCTCPMTPPCCPGCV  
LVLDGCGCKVCAKRLTESCHLHVCEPSQGLVCPAGPGHGAVCLIDEDGCEV  
NGRRLDETFKPNCRVLCRDGDGFTPLPCLSEDTPLMDCKPRKHIOVGKCP  
WVCDGVTPIQORSAOGHLSALVTPASADLPMNMTSTAMPCSTGCLGIATRVSN  
ONRFQGLIORLILCPRLCLARSHSSNNSAF"  
BASE COUNT 386 a 491 c 480 g 384 t  
ORIGIN  
Query Match 97.2%; Score 732.2; DB 10; Length 1741;  
Best Local Similarity 98.3%; Pred. No. 2e-168;  
Matches 740; Conservative 0; Mismatches 13; Indels 0; Gaps 0;  
1 ATGAGGGGAGCCCACTGATCATCTTGGCCACTTCCCTGCTTCATG 60  
|||||  
252 ATGAGGGGAGCCCACTGATCATCTTGGCCACTTCCCTGCTTCATG 321  
|||||  
61 GTGTGGCCAGCTGTGGCGGACCCGTGACCTGCTTGACACACCCAGTGGCA 120  
|||||  
322 GTGTGGCCAGCTGTGGCGGACCCGTGACCTGCTTGACACACCCAGTGGCA 381  
|||||  
121 CAGGGGTACCCCTGTGTGATGATGCTGTGGCTGTGTAAGTGTGACAGGAGCTG 180  
|||||  
382 CAGGGGTACCCCTGTGTGATGATGCTGTGGCTGTGTAAGTGTGACAGGAGCTG 441  
|||||  
181 GGGAGTCTCTGGACACCACTGATGTTCGACACCCAGCCAGGCGCTGTTGACGCT 240  
|||||  
442 ACGGAGTCTCTGGACACCACTGATGTTCGACACCCAGCCAGGCGCTGTTGACGCT 501  
|||||  
241 GGGGAGGAGCCCTGGGCGCATGGGCTGTGTCTCTTGGAGAGAGATGACGGTAGCTGT 300  
|||||  
502 GGGGAGGAGCCCTGGGCGCATGGGCTGTGTCTCTTGGAGAGAGATGACGGTAGCTGT 561  
|||||  
301 GAGGTGAATGGCGGAGGTACCTGGATGGAGAGACCTTTAAACCAATTGCAGGGTCTGT 360  
|||||  
562 GAGGTGAATGGCGGAGGTACCTGGATGGAGAGACCTTTAAACCAATTGCAGGGTCTGT 621  
|||||  
361 TGCCGCTGTGATGAGAGGTGGCTTACCTGCGCGCTGTGAGTGAAGATGTGGCGCTG 420  
|||||  
622 TGCCGCTGTGATGAGAGGTGGCTTACCTGCGCGCTGTGAGTGAAGATGTGAAGCTG 681  
|||||  
421 CCCAGCTGGAGCTGGCCAGGCGCCCAAGAGAAATACAGGTGGCGAAAGTGGCGCCGAG 480  
|||||  
682 CCCAGCTGGAGCTGGCCAGGCGCCCAAGAGAAATACAGGTGGCGAAAGTGGCGCCGAG 741  
|||||  
481 TGGGTATGTGACAGGAGGTGACACCGCGGATGACGCGCTCCACGCGCAGAGACACCA 540  
|||||  
742 TGGGTATGTGACAGGAGGTGACACCGCGGATGACGCGCTCCACGCGCAGAGACACCA 801  
|||||  
541 CTTTGTGCGCTTTGTATCTCTGCTGTGATGCTCTTGTCCAAATTGAGACACAGCC 600  
|||||

|||||  
Db 802 CTTTGTGCGCTTTGTATCTCTGCTGTGATGCTCTTGTGCGCAAAATTGAGACACAGCC 861  
|||||  
Qy 601 TGGGGCCCTGTGTCACACCACTGTGGGCTGGGATAGGACACCGAGTGTCCAAAGACAG 660  
|||||  
Db 862 TGGGGCCCTGTGTCACACCACTGTGGGCTGGGATAGGACACCGAGTGTCCAAAGACAG 921  
|||||  
Qy 661 CGATTCTGTGCAACTGAGATCCCAACGCGGCTGTGTGTCGCCAGACCTGTGCTGGACGC 720  
|||||  
Db 922 CGATTCTGTGCAACTGAGATCCCAACGCGGCTGTGTGTCGCCAGACCTGTGCTGGACGC 981  
|||||  
Qy 721 AGGAGCCACAGCTCATGTGACAGTCTTTCTTA 753  
|||||  
Db 982 AGGAGCCACAGCTCATGTGACAGTCTTTCTTA 1014  
|||||  
RESULT 2  
AR210324 1734 bp DNA linear PAT 20-JUN-2002  
LOCUS  
DEFINITION Sequence 17 from patent US 6387657.  
ACCESSION AR210324  
VERSION AR210324.1 GI:21512525  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1734)  
AUTHORS Borstein, D.A., Cohen, R.L., Goddard, A.D., Gurney, A.L., Hillan, K.J.,  
Lawrence, D.A., Levine, A.J., Pennica, D., Roy, M. Ann. and Wood, W.I.  
WISP polypeptides and nucleic acids encoding same  
Patent: US 6387657-A 17 14-MAY-2002;  
FEATURES  
source Location/Qualifiers  
1..1734  
/organism="unknown"  
BASE COUNT 355 a 491 c 495 g 393 t  
ORIGIN  
Query Match 87.5%; Score 659; DB 6; Length 1734;  
Best Local Similarity 93.0%; Pred. No. 1.6e-150;  
Matches 702; Conservative 0; Mismatches 50; Indels 3; Gaps 1;  
1 ATGAGGGGAGCCCACTGATCATCTTGGCCACTTCCCTGCTTCATG 60  
|||||  
257 ATGAGGGGAGCCCACTGATCATCTTGGCCACTTCCCTGCTTCATG 316  
|||||  
Qy 61 GTGTGGCCAGCTGTGGCGGACCCGTGACCTGCTTGACACACCCAGTGGCA 120  
|||||  
Db 317 GTGTGGCCAGCTGTGGCGGACCCGTGACCTGCTTGACACACCCAGTGGCA 376  
|||||  
Qy 121 CAGGGGTACCCCTGTGTGATGATGCTGTGGCTGTGTAAGTGTGACAGGAGCTG 180  
|||||  
Db 377 CCGGGGTACCCCTGTGTGATGATGCTGTGGCTGTGTAAGTGTGACAGGAGCTG 436  
|||||  
Qy 181 GGGGAGTCTCTGGACACCACTGATGTTCGACACCCAGCCAGGCGCTGTTGACGCT 240  
|||||  
Db 437 GGGGAGTCTCTGGACACCACTGATGTTCGACACCCAGCCAGGCGCTGTTGACGCT 496  
|||||  
Qy 241 GGGGAGGAGCCCTGGGCGCATGGGCTGTGTCTCTTGGAGAGAGATGACGGTAGCTGT 300  
|||||  
Db 497 GGGGAGGAGCCCTGGGCGCATGGGCTGTGTCTCTTGGAGAGAGATGACGGTAGCTGT 556  
|||||  
Qy 301 GAGGTGAATGGCGGAGGTACCTGGATGGAGAGACCTTTAAACCAATTGCAGGGTCTGT 360  
|||||  
Db 557 GAGGTGAATGGCGGAGGTACCTGGATGGAGAGACCTTTAAACCAATTGCAGGGTCTGT 616  
|||||  
Qy 361 TGCCGCTGTGATGAGAGGTGGCTTACCTGCGCGCTGTGAGTGAAGATGTGGCGCTG 420  
|||||  
Db 617 TGCCGCTGTGATGAGAGGTGGCTTACCTGCGCGCTGTGAGTGAAGATGTGGCGCTG 676  
|||||  
Qy 421 CCCAGCTGGAGCTGGCCAGGCGCCCAAGAGAAATACAGGTGGCGAAAGTGGCGCCGAG 480  
|||||  
Db 677 CCCAGCTGGAGCTGGCCAGGCGCCCAAGAGAAATACAGGTGGCGAAAGTGGCGCCGAG 736  
|||||







```

Db      10 ATGAGAGGACACCAAGACCACCTCCGCGCTTCTCCCTCTGCTCTCCTCAAG 69
Qy      61 GTGTGTGCGACAGCTGTGCGGACACCTGTACTCTGCTTGGACACCCAGTCCCA 120
Db      70 GTGCGTACCACTGTGCGCGGACACATGTACTGCTCCCTGCGACCTCCCGATGCCG 129
Qy      121 CAGGGGGATACCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
Db      130 CTGGAGATACCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 189
Qy      181 GGGGAGTCTGTGACCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 240
Db      190 GGGGAGCCCTGTGACCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 249
Qy      241 GGGGAGCCCTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
Db      250 GGGGAGGAGCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 309
Qy      301 GAGGTGAATGCGCGGAGTACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 360
Db      310 GAGGTGAATGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 369
Qy      361 TGCCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 420
Db      370 TGCCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 429
Qy      421 CCCAGCTGGAGTGGCCCAAGAGATATACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 480
Db      430 CCCAGCTGGAGTGGCCCAAGAGATATACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 489
Qy      481 TGGGTATGTGACCAAGAGATGACACCGCGGATCCAGCGCTTCAGCGCGGCGGAGACCAA 540
Db      490 TGGGTATGTGACCAAGAGATGACACCGCGGATCCAGCGCGCTTCAGCGCGGCGGAGACCAA 549
Qy      541 CTTTGTGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
Db      550 TTTTGTGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 609
Qy      601 TGGGCGCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660
Db      610 TGGGCGCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 669
Qy      661 CGATTCTGTGCAATGTGAGATCCAGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 720
Db      670 CGATTCTGTGCAATGTGAGATCCAGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 729
Qy      721 AGGAGCAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 752
Db      730 AGGAGCAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 761

```

```

RESULT 8
AF083500 1283 bp mRNA linear PRI 04-NOV-1998
LOCUS     Homo sapiens connective tissue growth factor-like protein
DEFINITION
ACCESSION AF083500
VERSION    AF083500.1 GI:3462835
KEYWORDS
SOURCE     Homo sapiens.
ORGANISM   Homo sapiens; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1283)
            Kumar,S., Hand,A.T., Connor,J.C., Dodds,R.A., Ryan,P.J.,
            Trill,J.G., Fisher,S.M., Siemmon,J.R., Lipschutz,D.B.,
            Bartholomew,V., James,I.E., Rieaman,D.J., Gowen,M. and Lee,J.C.
            Identification and cloning of CTGF-L from human osteoblasts, a
            novel cysteine rich protein containing an IGF binding domain
            Bone 23 (5), S240 (1998)
JOURNAL   Kumar,S.
REFERENCE 2 (bases 1 to 1283)
AUTHORS

```

```

TITLE     Direct Submission
JOURNAL   Submitted (11-AUG-1998) Bone & Cartilage Biology, UM 2109,
            Smithkline Beecham, 709 Swedeland Road, King of Prussia, PA 19406,
            USA
FEATURES
Source    Location/Qualifiers
          1..1283
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="20"
            /map="20q12-q13"
            /cell_type="primary osteoblast"
          9..761
            /note="CTGF-L; encodes IGF binding (IGFB), Von Willebrand
            factor type C (VWC) and thrombospondin type I (TSP1)
            domains; member of the CN (CTGF/cyrf61/nov) family; lacks
            the fourth carboxy-terminal domain present in other
            members of the CN family"
            /codon_start=1
            /product="connective tissue growth factor-like protein
            precursor"
            /protein_id="AAC70350.1"
            /db_xref="GI:3462836"
            /translation="MRGPKHLLAFSLILSLKYRQLCPPTCPMPRPPLGV
            LVLDCGCCRVCARRLGEPCDQHLVCDLSGLVCPGAPGGRALCLLAEDBSCEV
            NGRILREGEFTQPHCSIRKCEGDFVPLCSLEVRLLPMDCPHREVRVLGRCPPE
            WYCGGGGLGTPPLPAQGPQPSGLVSLPPGVPPEWSTAWGPESTTGGLMATRVSN
            ONRFRLETFORRLISRPCPPSRGRSPNSAF"
BASE COUNT 235 a 418 c 389 g 241 t
ORIGIN
Query Match 67.8%; Score 510.4; DB 9; Length 1283;
Best Local Similarity 79.9%; Pred. No. 3.3e-114;
Matches 601; Conservative 0; Mismatches 151; Indels 0; Gaps 0;
Qy      1 ATGAGGGGAGCCCACTGTATCCATCTTGCGGCACTCTGCTGCTTCTCAATG 60
Db      9 ATGAGAGGACACCAAGACCACCTCTGCGCTTCTCTCTCTGCTCTCTCAAG 68
Qy      61 GTGTGTGCGACAGCTGTGCGGACACCTGTACTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 120
Db      69 GTGCGTACCACTGTGTGCGCGGACACATGTACTGCTCCCTGCGACCTCCGAGTCCG 128
Qy      121 CAGGGGGATACCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
Db      129 CTGGAGATACCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 188
Qy      181 GGGGAGTCTGTGACCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 240
Db      189 GGGGAGCCCTGTGACCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 248
Qy      241 GGGGAGGCGCTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
Db      249 GGGGAGGACCCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 308
Qy      301 GAGGTGAATGCGCGGAGTACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 360
Db      309 GAGGTGAATGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 368
Qy      361 TGCCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 420
Db      369 TGCCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 428
Qy      421 CCCAGCTGGAGTGGCCCAAGAGATATACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 480
Db      429 CCCAGCTGGAGTGGCCCAAGAGATATACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 488
Qy      481 TGGGTATGTGACCAAGAGATGACACCGCGGATCCAGCGCTTCAGCGCGGCGGAGACCAA 540
Db      489 TGGGTATGTGACCAAGAGATGACACCGCGGATCCAGCGCGCTTCAGCGCGGCGGAGACCAA 548
Qy      541 CTTTGTGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
Db      549 TTTTGTGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 608

```



|            |  |   |     |        |                 |
|------------|--|---|-----|--------|-----------------|
| Qy         | 601  | TGGGGCCCCCTCTCAACCAACCTGGGGGCTGGGGATAGCCACCGAGTGTCCAACGAGAC       | 660 |        |                 |
| Db         | 609  | TGGGAGACCTTCTCTGCAGCCACTCTGTGGCTGGGGCATGATGGCCACCGGGTGTCCAACGAGAC | 668 |        |                 |
| Oy         | 661  | CGATTCTTGCCAACTGGAGATCCAAACGGCCGCTGTGTCTGCCCCAGACCCCTGCTTGGCAACC  | 720 |        |                 |
| Db         | 669  | CGCTTCTCCCACTGGAGAACCCACGCGCGCTGTGTCTGTCAGAGGCCCTTGGCCACCCCTCC    | 728 |        |                 |
| Oy         | 721  | AGGAGCCACGACTCATGGACAGAGCTCTTCA                                   | 752 |        |                 |
| Db         | 729  | AGGGTCCGACTCTCCACAAACAGTGCCTTCA                                   | 760 |        |                 |
| RESULT 9   | AR210322   | 1293 bp   | DNA | Linear | PAT 20-JUN-2002 |
| LOCUS      | AR210322   |   |     |        |                 |
| DEFINITION | Sequence 13 from patent US 6387657.                                |   |     |        |                 |
| ACCESSION  | AR210322   |   |     |        |                 |
| VERSION    | AR210322.1   | GI:21512523   |     |        |                 |
| KEYWORDS   |  |   |     |        |                 |
| SOURCE     | Unknown.   |   |     |        |                 |
| ORGANISM   | Unknown.   |   |     |        |                 |
| REFERENCE  | Unclassified.  |   |     |        |                 |
| AUTHORS    | 1 (bases 1 to 1293)  |   |     |        |                 |
| TITLE      | Boistein,D.A., Cohen,R.L., Goddard,A.D., Gurney,A.L., Hillan,K.J., |   |     |        |                 |
| JOURNAL    | Lawrence,D.A., Levine,A.J., Pennica,D., Roy,M.Ann. and Wood,W.I.   |   |     |        |                 |
| FEATURES   | WISP polypeptides and nucleic acids encoding same                  |   |     |        |                 |
| source     | Patent: US 6387657-A 13 14-MAY-2002;                               |   |     |        |                 |
|            | Location/Qualifiers  |   |     |        |                 |
|            | 1..1293  |   |     |        |                 |
|            | /organism="unknown"  |   |     |        |                 |
| BASE COUNT | 232 a 425 c 393 g 243 t  |   |     |        |                 |
| ORIGIN     |  |   |     |        |                 |

| Query Match           | 67.8% | Score 510.4  | DB 6           | Length 1293 |
|-----------------------|-------|--|----------------|-------------|
| Best Local Similarity | 79.9% | Pred. No. 3.3e-114   |                |             |
| Matches               | 601   | Conservative 0   | Mismatches 151 | Indels 0    |
| QY                    | 1     | ATGAGGGGAGCCACATGATTCATCTTGTGGCCACTTCTCTCTCTCTCTCTCTCTCATG     | 60             |             |
| Db                    | 22    | ATGAGAGGACCCAGAGACCACCTCTGTGGCTCTCTCTCTCTCTCTCTCTCTCAAG        | 81             |             |
| QY                    | 61    | GTGTGTGCCACGTGTGTGCCGAGACACCTGTACTGTCTTGTGGACACCCCAATGCCA      | 120            |             |
| Db                    | 82    | GTGTGTGCCACGTGTGTGCCGAGACACCTGTACTGTCTTGTGGACACCCCAATGCCG      | 141            |             |
| QY                    | 121   | CAGGGGTATACCCCTGTGTCTGTGAATGGCTGTGGCTGTAAAGTGTGCACGAGAGCTG     | 180            |             |
| Db                    | 142   | CTGGGAGTATACCCCTGTGTCTGTGAATGGCTGTGGCTGTAAAGTGTGCACGAGAGCTG    | 201            |             |
| QY                    | 181   | GGGAGTCTCTGCGACACCTGCTATGTCTGCGACCCCGACAGGGCTGTGTGACGCT        | 240            |             |
| Db                    | 202   | GGGAGGCTCTGCGACACCTGCTATGTCTGCGACCCCGACAGGGCTGTGTGACGCT        | 261            |             |
| QY                    | 241   | GGGGAGGAGCCCTGTGGGCGCATGGGGCTGTGTCTCTTGATGTGAAGATGAGCGTAGCT    | 300            |             |
| Db                    | 262   | GGGGAGGAGCCCTGTGGGCGCATGGGGCTGTGTCTCTTGATGTGAAGATGAGCGTAGCT    | 321            |             |
| QY                    | 301   | GAGGTGAATGGCCCGAGGTACTGTGTGATGGAGAGACCTTTAAACCCAAATTGACGGATCTG | 360            |             |
| Db                    | 322   | GAGGTGAATGGCCCGCTGTGTATGTGGAGAGGGAGACCTTCAAGCCCACTGACGATCCGC   | 381            |             |
| QY                    | 361   | TGCGCGTGTGAATGGAGGTGTTCACCTGCGCGCGGTGCACGTGAGAGATGTCCGCTG      | 420            |             |
| Db                    | 382   | TGCGCGTGTGAATGGAGGTGTTCACCTGCGCGCGGTGCACGTGAGAGATGTCCGCTG      | 441            |             |
| QY                    | 421   | CCGAGCTGGAGCTGCCACGCGCCCAAGAGAAATTACAGGTGCCAGGAAAGTGTGCCCGAG   | 480            |             |
| Db                    | 442   | CCGAGCTGGAGCTGCCCGCCCAAGAGAAATTACAGGTGCCAGGAAAGTGTGCCCGAG      | 501            |             |
| QY                    | 481   | TGGGTATGTGACCGAGGAGTACACCGGCGATCCAGCGCTCCACGCGCGCAAGAACCA      | 540            |             |

|    |     |  |   |     |
|----|-----|--|---|-----|
| Db | 502 | TGGGCTGTGGGACC   | AAAGGAGGGGAGATGTGGGAGCCAGCCCTTCCAGCCCAAGAACCCAG | 561 |
| OY | 541 | CTTTCTGGCCCTTGTCACCTCCCTGCTCTGTCTATGTCTCTTGTCCAAATTTGGAGCAGCAGCC     | 600   |     |
| Db | 562 | TTTTCTGGCCTGTCTCTCTTCTCCCTGCCCCCTGATGTCTCCCTCCCAAGAAATGAGCAGCGCC     | 621   |     |
| OY | 601 | TGGGGCCCTCTCTCAACCAACCTGTGTGGGCTGGGGCATAGCAGCAGCCGAGTGTCCACAGAGAAC   | 660   |     |
| Db | 622 | TGGGAGACCTCTCTGTGACCAACCTGTGTGGGCTGGGGCATAGCAGCAGCCGAGTGTCCACAGAGAAC | 681   |     |
| OY | 661 | CGATTCTTCCCACTTGGAGATCCAAAGCCGCTGTGTCTGTCCAGAGACCTGCTTGGCAGCC        | 720   |     |
| Db | 682 | CGCTCTTGCACACTGGAGAGCAGCAGCGCCGCTGTGTCCAGAGCCCTGTGCCAACCCCTTCC       | 741   |     |
| OY | 721 | AGGAGCCCACTCTCATGGAGCAGTGTCTTCTA                                     | 752   |     |
| Db | 742 | AGGGGTGCAGTCTCCACAAACAGTGTCTTCTA                                     | 773   |     |

| RESULT                | 10  | 1293 bp     | DNA | linear | PAT 20-JUN-2002 |
|-----------------------|---|-------------|-----|--------|-----------------|
| LOCUS                 | AR210323/C  |             |     |        |                 |
| DEFINITION            | Sequence 14 from patent US 6387657.   |             |     |        |                 |
| ACCESSION             | AR210323  |             |     |        |                 |
| VERSION               | AR210323.1  | GI:21512524 |     |        |                 |
| KEYWORDS              |   |             |     |        |                 |
| SOURCE                | Unknown.  |             |     |        |                 |
| ORGANISM              | Unknown.  |             |     |        |                 |
| REFERENCE             | 1 (bases 1 to 1293)   |             |     |        |                 |
| AUTHORS               | Botstein,D.A., Cohen,R.L., Goddard,A.D., Gurney,A.L., Hillan,K.J., Lawrence,D.A., Levine,A.J., Pennica,D., Roy,M.Ann. and Wood,W.I. |             |     |        |                 |
| TITLE                 | WISP polypeptides and nucleic acids encoding same   |             |     |        |                 |
| JOURNAL               | Patent: US 6387657-A 14 14-MAY-2002;  |             |     |        |                 |
| FEATURES              | Location/Qualifiers   |             |     |        |                 |
| source                | 1..1293   |             |     |        |                 |
| BASE COUNT            | 243 a 393 c 425 g 232 t   |             |     |        |                 |
| ORIGIN                | /organism="unknown"   |             |     |        |                 |
| Query Match           | 67.8%; Score 510.4; DB 6; Length 1293;  |             |     |        |                 |
| Best Local Similarity | 79.9%; Pred. No.3.3e-114;   |             |     |        |                 |
| Matches               | 601; Conservative 0; Mismatches 151; Indels 0; Gaps 0;  |             |     |        |                 |
| QY                    | 1 ATGAGGGGAGCCCACTGATCTCATCTTCTGCGCCACTTCTCTCTCTCTCTCAATG 60  |             |     |        |                 |
| Db                    | 1272 ATGAGAGGAGCAACACGAAAGACCACCTCTGCGCTCTCTCTCTCTCTCTCTCAAG 1213   |             |     |        |                 |
| QY                    | 61 GTGTGTGCCCAAGCTGTGCGCGGAGAACCCCTGTACCTGTCTCTTGTGAGACACACCCAGTGC 120  |             |     |        |                 |
| Db                    | 1212 GTGCTTACCCAGCTGTGTGCGCGAGACATGTACTCTTCCCTGCGCCACTCCCGATGCCG 1153   |             |     |        |                 |
| QY                    | 121 CAGGGGTACCCCTGTGTGCTGTGATGGAGCTGTGTGCTGTAAAGTGTGTGACAGAGCTG 180   |             |     |        |                 |
| Db                    | 1152 CTGGGAGTACCCCTGTGTGCTGTGATGGAGCTGTGTGCTGTCTCCGGGTATGTGTGACAGGCGCTG 1093  |             |     |        |                 |
| QY                    | 181 GGGGAGTCTCTGAGACACCTGATGTGTGCAACCCCAAGCCGAGGCGCTGTTTGTAGCCT 240   |             |     |        |                 |
| Db                    | 1092 GGGGAGCCCTGAGACACCTGACAGTGTGCGAGCGCACGAGGCGCTGTGTGCGCAGGCC 1033  |             |     |        |                 |
| QY                    | 241 GGGGAGGAGCCCTGCGGCGCATGAGGGGCTGTGTGTCTCTTGTGATGAGAGATGACGATAGCTGT 300   |             |     |        |                 |
| Db                    | 1032 GGGGAGGAGCCCGGTGGGCGGGGCGCCGTGTGCTCTTGTGGAGAGGAGGAGAGCGAGCTGT 973  |             |     |        |                 |
| QY                    | 301 GAGGTGATGTGCGCGAGGTACTCTGGATGTGAGAGACCTTAAACCAATTTGAGAGGCTCTG 360   |             |     |        |                 |
| Db                    | 972 GAGGTGAGCGGCGCTGTATGTGGGAGAGGAGACCTTCCAGCCCACTGTGAGCATCTCCG 913   |             |     |        |                 |
| QY                    | 361 TGCCGCTGTGTGAGAGGAGGCTTACCTGCGCTGTGCGGCTGTGACAGTGTGAGATGTGCGGCTG 420  |             |     |        |                 |
| Db                    | 912 TGCCGCTGTGAGAGGAGGCGGCTTACCTGCGCTGTGCGGCTGTGACAGGAGATGTGCGGCTG 853  |             |     |        |                 |
| QY                    | 421 CCCAGCTGTGAGCTGCCCAAGCCCAAGAGATACAGGTGTCAGAGAAAGTGTGCTGCCCGAG 480   |             |     |        |                 |



[illegible]

| QY         | 481        | TGGGTATGTGACACGAGGAGTGACACGCGCGATCCAGGCGTCCACAGGGGCAAGACACCA      | 540                         |
|------------|------------|---|-----------------------------|
| Db         | 636        | TGGGTGTGCGGCCACAGAGGAGGGGACGTGGGAGCCAGCCCTTTCACAGCCCAAGAGACCCAG   | 695                         |
| QY         | 541        | CTTTCGTGCGCCCTTGTGACCTCTGCTCTGCTGTGTATGCTCTGTGTCCAAATTTGAGCACAGCC | 600                         |
| Db         | 696        | TTTTCTTGCGCCTGTCTCTTCCCTGTCGCCCTGCTGTCCCTGCGCCAGATGTGAGACAGGGCC   | 755                         |
| QY         | 601        | TGGGGCCCTGCTGCTACCCACCTGTGGGCTGGGGCATATGACACCGAGTGTCCACACAGAAC    | 660                         |
| Db         | 756        | TGGGGACCCCTGCTGACACCTGTGGGCTGGGGCATACCCCGGTGTCCACACAGAAC          | 815                         |
| QY         | 661        | CGATCTTGCCACATCGAGATCCACAGGCGCGCTGTGTGTGGCCACAGCCCTGCTGGAGCC      | 720                         |
| Db         | 816        | CGCTCTTGCGCAGCTGGAGAGACCCAGCGCGCTGTGCTGTGTCCAGGCGCTGCCACCTCC      | 875                         |
| QY         | 721        | AGGAGCCACAGCTCATGTGAAACAGTCTTTCCTA                                | 752                         |
| Db         | 876        | AGGGGTGCGCAGTCCACAAACAGTCCCTCTA                                   | 907                         |
| RESULT 13  |            |   |                             |
| LOCUS      | BC017782   | 1450 bp   | mRNA linear PRI 06-DEC-2001 |
| DEFINITION |            |   |                             |
| ACCESSION  | BC017782   |   |                             |
| VERSION    | BC017782.1 |   |                             |
| KEYWORDS   |            |   |                             |
| SOURCE     |            |   |                             |
| ORGANISM   |            |   |                             |
| REFERENCE  |            |   |                             |
| AUTHORS    |            |   |                             |
| TITLE      |            |   |                             |
| JOURNAL    |            |   |                             |
| REMARK     |            |   |                             |
| COMMENT    |            |   |                             |

```

BASE COUNT      272 a      457 c      441 g      280 t
ORIGIN
1  /Installation="MRGTPKHHLAFLSLCLTSVAPOLCTPTCTCMPPRCCLGVA
2  /LUDGGCCGRCVARGRGLSPOLHVACASOGVYCPGAGRGALCLLADDSCEV
3  NGRLYRGEFTGPOPLSRCEGDEGFCVPLGSDVILPMDCHPRVETKCCCE
4  WGGGGGGLGPOPLPAGPOSGVYSLPVYCEPEMSTANGPCTTCGLMATRYNSN
5  QNRCLRLTORLRLSLRCPSPRRSRNSNF"

```

|                           |        |                      |           |              |
|---------------------------|--------|----------------------|-----------|--------------|
| Query Match:              | 67.88% | Score 510.4:         | DB 9      | Length 1450: |
| Best Local Similarity:    | 79.98% | -Pred. No. 3.2e-114: |           |              |
| Matches 601, Conservative | 0:     | Mismatches 151:      | Indels 0: | Gaps 0:      |

|    |     |  |     |
|----|-----|--|-----|
| OY | 1   | ATGAGGGGACCCACATGATCCATCTTGTGGCACTGTCATCTGCTCTTCAGAG               | 60  |
| Db | 165 | ATGAGAGGACACACCGAAGACCACCTCTGGGCTTTCTCCCTCTCTGCTCTTCAGAG           | 224 |
| OY | 61  | GTCGTGTCGCCAGCTGTGTGGGACACCCCTGTACTGTCTCTTGGAGACACCCAGTGGCCA       | 120 |
| Db | 225 | TGCGGTACCCAGCTGTGTGGGACACAGATGTACTGTGCCCTGGCCACCTCCCCGAGTGGCCG     | 284 |
| OY | 121 | CAGGGGGTACCCCTGTGTGTGGATGTGTGTGGTGTGTGTAAAGTGTGTGACAGGAGCTG        | 180 |
| Db | 285 | CTGGGAGTACCCCTGTGTGTGGATGTGTGTGGTGTGTGTGCGGGGTATGTGACAGGGGGCTG     | 344 |
| OY | 181 | GGGGAGTCTCTGACACACCTGATGTGTGGACCCAGCCGAGGGGGCTGTGTGTGACGCT         | 240 |
| Db | 345 | GGGGAGCCCTCTGACACACCTGATGTGTGGACCCAGCCGAGGGGGCTGTGTGTGACGCT        | 404 |
| OY | 241 | GGGGCAGGCCCTGGCGGGCATGTGGGGCTGTGTCTTGTGATGTAGAGATGACGATAGCTGT      | 300 |
| Db | 405 | GGGGCAGGACCCGGGTGGCGGGGGGGCCCTGTGTCTTGTGCGAGAGGACGACGACGCTGT       | 464 |
| OY | 301 | GAGGTGAATGGCCCGCAGGTACTGTGGATGTGAGACACTTTAAACCAATTGCAAGGCTCTG      | 360 |
| Db | 465 | GAGGTGAACGGCCCGCCTGTATCTGGGGAAGGGGAGACCTTTCAGCCCACTGCAGCATCCGC     | 524 |
| OY | 361 | TGCGCGTGTGATGAGCGGTGTACCTGACCGCTGGCGGCTGTGACAGATGTGGGGCTG          | 420 |
| Db | 525 | TGCGCGTGTGAGAGACGGCGCTTACCTGTGCTGTGCGGCTGTGTGAGAGAGATGTGGGGCTG     | 584 |
| OY | 421 | CCGAGCTGTGGACTGCCACGCCCCAGAGAAATACAGGTGCCAGGAAATGCTGCCCCAG         | 480 |
| Db | 585 | CCGAGCTGTGGACTGCCACGCCCCAGAGAGGAGTTCAGAGTTCGAGGGAATGTGCTGCCCTAG    | 644 |
| OY | 481 | TGGGTATGTACCAAGGAGATGACACCGGCGATCCAGCGCTTCACGGCGCAAGACCA           | 540 |
| Db | 645 | TGGGTATGTGCGCCCAAGGAGGGGAGACTGTGGGACCCAGGCCCTTCACAGCCCAAGGAGACCCAG | 704 |
| OY | 541 | CTTTCTGCGCCCTGTGACACTGCTGCGCTGTGCTGATGCTGCTTGTCCAAATTTGAGACAGCC    | 600 |
| Db | 705 | TTTTCTGCGCTGTGCTGCTTCTCCCTGCCCTGTGTGCTGCTGCTCCCAAGATGAGACAGCCG     | 764 |
| OY | 601 | TGGGGCCCTGTGCTGACACCTGTGTGGGTGGGATAGGACCGAGTGTCCAAACAGAAC          | 660 |
| Db | 765 | TGGGGACCCCTGTGTGACACCTGTGTGGGTGGGATAGGACCGAGTGTCCAAACAGAAC         | 824 |
| OY | 661 | CGATTCTGCAACTGTGAGATCCAAAGCCGCGCTGTGTGTGCGGAGACCTGTGCTGGACCC       | 720 |
| Db | 825 | CGATTCTGCGCACTGTGAGATCCCAAGCCGCGCTGTGTGTGCGGAGACCTGTGCTGGACCC      | 884 |
| OY | 721 | AGGAGCCACAGCTCATGGAACAGTCTTCA 752                                  |     |
| Db | 885 | AGGGGTGCAATCCACAAAGTGTCTTCA 916                                    |     |

|            |                                     |
|------------|-------------------------------------|
| RESULT     | 14                                  |
| LOCUS      | AR210337                            |
| DEFINITION | Sequence 38 from patent US 6387657. |
| ACCESSION  | AR210337                            |
| VERSION    | AR210337.1                          |
| KEYWORDS   | GI:2512542                          |
| SOURCE     | Unknown.                            |
|            | linear PAT 20-JUN-2002              |

| ORGANISM   | Unknown.  |
|------------|---|
| REFERENCE  | Unclassified.   |
| AUTHORS    | 1 (bases 1 to 738)  |
| TITLE      | Botstein,D.A., Cohen,R.L., Goddard,A.D., Gurney,A.L., Hillan,K.J., Lawrence,D.A., Levine,A.J., Pennica,D., Roy,M.Ann. and Wood,W.I. |
| JOURNAL    | WISP polypeptides and nucleic acids encoding same   |
| FEATURES   | Patent: US 6387657-A 38 14-MAY-2002;  |
| source     | Location/Qualifiers   |
|            | 1..738  |
| BASE COUNT | /organism="unknown" 104 a 272 c 238 g 124 t   |
| ORIGIN     |   |

|    |                       |  |                     |       |                 |                   |
|----|-----------------------|--|---------------------|-------|-----------------|-------------------|
|    | Query Match           | 66.6%;   | Score 501.2;        | DB 6; | Length 738;     |                   |
|    | Best Local Similarity | 79.9%;   | Pred. No. 6-2e-112; |       |                 |                   |
|    | Matches               | 590;   | Conservative        | 0;    | Mismatches 148; | Indels 0; Gaps 0; |
| OY | 13                    | CCACGTATCATCTTGTGGCACCATTCTTCCTTGCCCTTCTCTCAATGAGTGTGCCAG          | 72                  |       |                 |                   |
| Dd | 1                     | CCGAAGAACCACACTCTCTGGCCCTTCTCCCTCTCTCTCTCTCAAAAGGTGGTAACCG         | 60                  |       |                 |                   |
| OY | 73                    | CTGAGCGGACAACCTGTACCTGACCTGTGGACACCAACCCAGCCAGTGCACAGGGGGTACC      | 132                 |       |                 |                   |
| Dd | 61                    | CTGTGCCCGAACCATTGACTCTGGCCCCCTGGCACACTCCCAGATGCCCTGGGATATCCC       | 120                 |       |                 |                   |
| OY | 133                   | CTGTGTCTGATGTGGCTGTGGCTGTCTGTAAAGTGTGTGCACGAGAGGCTGGGGAGTCTGC      | 192                 |       |                 |                   |
| Dd | 121                   | CTGGTGCTGATATGGCTGTGGCTGTGGCTGTGGCGGGATGTGCACGGCGGCTGGGGAGCCCTGC   | 180                 |       |                 |                   |
| OY | 193                   | GACCACCTGCATGTCTGCGACCCCAACGACGAGGGCTGTGTTGTCAAGCTTGGGGCAGCCCT     | 252                 |       |                 |                   |
| Dd | 181                   | GACCAACTCCACGTCCTGGACGCCACGACGAGGCTGTGTCTGCACCCCGGGGCAGAGCCC       | 240                 |       |                 |                   |
| OY | 253                   | GGCGGCACTGGAGGCGTGTGTCTCTCTTGATGAGATGACGGTAGTGTGAGGTGAATGAC        | 312                 |       |                 |                   |
| Dd | 241                   | GGTGGCGCGGGGGGCCCTGTGCTCTGTGGCAGAGAACGACAGCAGTGTGAGGTGAACGCG       | 300                 |       |                 |                   |
| OY | 313                   | CGCAGTACCTGTGATGAGAGACCTTTAAACCCAATTGCAGGGTCTGTGCCGCTGTGAT         | 372                 |       |                 |                   |
| Dd | 301                   | CGCCTGTATCGGGAAGGGGAAACCTTCACAGCCCTCACACATCCCTGCGCGTGCAG           | 360                 |       |                 |                   |
| OY | 373                   | GACGGTGGCTTACCTGTGCTCCGCTGGACGTGAGATGTGGGCTGCCACGAGGGAG            | 432                 |       |                 |                   |
| Dd | 361                   | GACGGCGGCTTACCTGTGCTCCGCTGTGACGACGAGATGTGGGCTGCCACGCTGGGAC         | 420                 |       |                 |                   |
| OY | 433                   | TGCCACACGCCCAAGAGATAACAGTGTGCAGGAAAGTCTGCCCCGAGTGGGTATGTAC         | 492                 |       |                 |                   |
| Dd | 421                   | TGCCCCACCCCAAGAGGTGTGAAGTCTGTGGGCAAGTCTGCCCCGAGTGGGTATGTAC         | 480                 |       |                 |                   |
| OY | 493                   | CAGGAGTGCACACGGGCGATCAGCGCTTCACAGGCGCAAGACACCAATTCTTGCCCTT         | 552                 |       |                 |                   |
| Dd | 481                   | CAAAGAGGGGACATGGGGACCCAGGCCCTTCACAGGCCCAAGGACCCCAATTCTTGCCCTT      | 540                 |       |                 |                   |
| OY | 553                   | GTCATCCCTGTGCTGTGATGTCTTCCTGTCCAAATTTGGAGACAGCTGGGGGACCTTGC        | 612                 |       |                 |                   |
| Dd | 541                   | GTCCTCTTCCCTGCCCCCTGTGTGTCTCCCTGTGCCCAAGATGAGACAGCGGCTTGGGAGCCCTGC | 600                 |       |                 |                   |
| OY | 613                   | TCAACACACTTGGGGTGGGATAGCCACCCGAGTGTCCAAACAGAACCGATTCTGCCAA         | 672                 |       |                 |                   |
| Dd | 601                   | TCGACACACTTGGGGTGGGATAGCCACCCGAGTGTCCAAACAGAACCGCTTCTGCCGA         | 660                 |       |                 |                   |
| OY | 673                   | CTGAGATCCCAACGCGGCTGTGTCTGTGCCACAGCCCTGTGCGACCCAGAGCCACAGC         | 732                 |       |                 |                   |
| Dd | 661                   | CTGGAGACCCACGCGCGCTGTGTCTGTGTCCAGGCGCTGCCACCTCCAGGGGTGCAGT         | 720                 |       |                 |                   |
| OY | 733                   | TCATGGAACAGTGGCTTC 750   |                     |       |                 |                   |
| Dd | 721                   | CCACAAAACATGCTTTC 738  |                     |       |                 |                   |

|          |                                   |
|----------|-----------------------------------|
| RESULT   | 15                                |
| AR210338 |                                   |
| LOCUS    |                                   |
| AR210338 | 841 bp DNA linear PAT 20-JUN-2007 |





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 23, 2003, 09:03:08 ; Search time 199.398 Seconds  
(without alignments)  
8504.365 Million cell updates/sec

Title: US-10-010-408-3  
Perfect score: 753  
Sequence: 1 ATGAGGGGACAGCCACTGAT.....CATGAACAGTCTTCTTAA 753

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues  
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_101002.\*  
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*  
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*  
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*  
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*  
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*  
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*  
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*  
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*  
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*  
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*  
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*  
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*  
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*  
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*  
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*  
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*  
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*  
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*  
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*  
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*  
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*  
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*  
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*  
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1          | 753   | 100.0       | 753    | 20    | AAZ07517    |
| 2          | 753   | 100.0       | 1708   | 20    | AAZ07516    |
| 3          | 681   | 90.4        | 681    | 20    | AAZ07521    |
| 4          | 659   | 87.5        | 1734   | 20    | AAZ07521    |
| 5          | 657   | 87.3        | 753    | 20    | AAZ07521    |
| 6          | 657   | 87.3        | 1257   | 20    | AAZ07521    |
| 7          | 657   | 87.3        | 1257   | 20    | AAZ07521    |
| 8          | 657   | 87.3        | 1257   | 20    | AAZ07521    |
| 9          | 657   | 87.3        | 1257   | 20    | AAZ07521    |

|    |       |      |       |    |          |                      |
|----|-------|------|-------|----|----------|----------------------|
| 10 | 510.4 | 67.8 | 1266  | 22 | AAZ07517 | Human angiogenesis   |
| 11 | 510.4 | 67.8 | 1285  | 19 | AAZ07517 | Human connective t   |
| 12 | 510.4 | 67.8 | 1293  | 20 | AAZ07517 | Human WISP-2 prote   |
| 13 | 510.4 | 67.8 | 1309  | 22 | AAZ07517 | Connective tissue    |
| 14 | 510.4 | 67.8 | 1522  | 20 | AAZ07517 | Human growth facto   |
| 15 | 510   | 67.7 | 1337  | 22 | AAZ07517 | Human secreted pro   |
| 16 | 508.8 | 67.6 | 1352  | 20 | AAZ07517 | Human secreted pro   |
| 17 | 508.4 | 67.5 | 750   | 20 | AAZ07517 | Human WISP-2 prote   |
| 18 | 501.2 | 66.6 | 738   | 20 | AAZ07517 | Human WISP-2 prote   |
| 19 | 499.4 | 66.3 | 1267  | 21 | AAZ07517 | Human PRO261 CDNA    |
| 20 | 496.4 | 65.9 | 841   | 20 | AAZ07517 | Human WISP-2 prote   |
| 21 | 210   | 27.9 | 210   | 20 | AAZ07517 | Rat HICP IGFBP dom   |
| 22 | 199   | 26.4 | 2136  | 22 | AAZ07517 | Rat HICP full-length |
| 23 | 177   | 23.5 | 177   | 20 | AAZ07517 | Rat HICP WVC domain  |
| 24 | 174   | 23.1 | 174   | 20 | AAZ07517 | Rat HICP WVC domain  |
| 25 | 169.8 | 22.5 | 586   | 22 | AAZ07517 | Human foetal liver   |
| 26 | 169.8 | 22.5 | 586   | 22 | AAZ07517 | Probe #698 for ge    |
| 27 | 169.8 | 22.5 | 586   | 22 | AAZ07517 | Human brain expres   |
| 28 | 169.8 | 22.5 | 586   | 22 | AAZ07517 | Human bone marrow    |
| 29 | 169.8 | 22.5 | 586   | 22 | AAZ07517 | Probe #6485 for ge   |
| 30 | 169.8 | 22.5 | 586   | 22 | AAZ07517 | Probe #8254 used t   |
| 31 | 169.8 | 22.5 | 586   | 22 | AAZ07517 | Human genome-deriv   |
| 32 | 169.8 | 22.5 | 13255 | 22 | AAZ07517 | Human immune/haema   |
| 33 | 163   | 21.6 | 2075  | 16 | AAZ07517 | Connective tissue    |
| 34 | 163   | 21.6 | 2075  | 18 | AAZ07517 | Human connective t   |
| 35 | 163   | 21.6 | 2075  | 19 | AAZ07517 | Connective tissue    |
| 36 | 163   | 21.6 | 2075  | 19 | AAZ07517 | Human connective t   |
| 37 | 163   | 21.6 | 2075  | 20 | AAZ07517 | Human connective t   |
| 38 | 163   | 21.6 | 2075  | 21 | AAZ07517 | Human connective t   |
| 39 | 163   | 21.6 | 2075  | 21 | AAZ07517 | Human connective t   |
| 40 | 163   | 21.6 | 2075  | 22 | AAZ07517 | Human connective t   |
| 41 | 163   | 21.6 | 2075  | 22 | AAZ07517 | Human connective t   |
| 42 | 163   | 21.6 | 2075  | 22 | AAZ07517 | Human connective t   |
| 43 | 163   | 21.6 | 2075  | 22 | AAZ07517 | Human connective t   |
| 44 | 163   | 21.6 | 2075  | 22 | AAZ07517 | Human benign prost   |
| 45 | 163   | 21.6 | 2312  | 22 | AAZ07517 | Nucleotide sequenc   |
|    |       |      |       |    |          | Human shear stress   |

## ALIGNMENTS

|          |  |                         |
|----------|--|-------------------------|
| RESULT 1 | AAZ07517   | standard; cDNA; 753 BP. |
| XX       | AAZ07517   |                         |
| AC       | AAZ07517   |                         |
| XX       | AAZ07517   |                         |
| DT       | 26-NOV-1999  | (first entry)           |
| XX       | 26-NOV-1999  |                         |
| DE       | Rat HICP polypeptide coding sequence.                                    |                         |
| XX       | Rat HICP polypeptide coding sequence.                                    |                         |
| KW       | Heparin-induced CCN-like protein; HICP; cell-associated activity; sg;    |                         |
| KW       | cardiovascular disorder; aberrant cell proliferation; fibrotic disorder. |                         |
| XX       |  |                         |
| OS       | Rattus sp.   |                         |
| XX       |  |                         |
| PN       | W09947556-A2.  |                         |
| XX       | W09947556-A2.  |                         |
| PD       | 23-SEP-1999.   |                         |
| XX       | 23-SEP-1999.   |                         |
| PF       | 18-MAR-1999;   | 99WO-US05999.           |
| XX       | 18-MAR-1999;   |                         |
| PR       | 19-MAR-1998;   | 98US-0044273.           |
| XX       | 19-MAR-1998;   |                         |
| PA       | (TUFT ) TUFTS COLLEGE.   |                         |
| XX       | (TUFT ) TUFTS COLLEGE.   |                         |
| PI       | Castellot JJ;  |                         |
| XX       | Castellot JJ;  |                         |
| DR       | WPI: 1999-562060/47.   |                         |
| XX       | WPI: 1999-562060/47.   |                         |
| DR       | P-PSDB: AAZ07517.  |                         |
| XX       | P-PSDB: AAZ07517.  |                         |
| PT       | Nucleic acid sequences encoding rat heparin-induced CCN-like protein,    |                         |
| XX       | used in methods to identify modulators or in diagnostic applications     |                         |







KW tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;  
 KW kidney disorder; bone-related disorder; osteoporosis; trauma; burn;  
 KW connective tissue disorder; catabolic state; inflammation;  
 KW testicular-related disorder; angiogenesis; immunological disorder; ss.  
 OS Mus sp.  
 XX W09921998-A1.  
 XX PD 06-MAY-1999.  
 XX PE 29-OCT-1998; 98WO-US22991.  
 XX PR 14-APR-1998; 98US-0081695.  
 XX PR 29-OCT-1997; 97US-0063704.  
 XX PR 03-FEB-1998; 98US-0073612.  
 XX PA (GENETECH ) GENENTECH INC.  
 XX PI Botstein DA, Cohen RL, Goddard A, Gurney AL, Hillan K;  
 XX PI Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WI;  
 DR WPI: 1999-337420/28.  
 DR P-PSDB: AAY17651.  
 XX PT New isolated wnt-1 induced secreted polypeptides, WISP-1, 2 and 3  
 XX PS Example 2; Page 178-179; 284pp; English.  
 XX CC The present invention describes wnt-1 induced secreted polypeptides,  
 CC WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2  
 CC and WISP-3 have homology to connective tissue growth factor (CTGF).  
 CC Products from the present invention can be used to treat WISP-related  
 CC disorders such as breast, ovarian, and colon cancer or melanoma. The  
 CC products can be used to treat arteriosclerosis. The products can also be  
 CC used to treat other diseases e.g. benign and malignant tumours,  
 CC leukemia and lymphoid malignancies, neuronal, glial, astrocytal,  
 CC hypochalamic and other glandular, macrophagal, epithelial, stromal, and  
 CC blastocelamic disorders, haematopiasis-related disorders, tissue-growth  
 CC disorders, skin disorders, desmoplasia, fibrotic lesions, kidney  
 CC disorders, bone-related disorders such as osteoporosis, trauma such as  
 CC burns, incisions, and other wounds, connective tissue disorders,  
 CC catabolic states, testicular-related disorders, and inflammation. The  
 CC antigenic and immunologic disorders including arteriosclerosis. The  
 CC products can also be used for detection and diagnosis especially of  
 CC individuals with neoplastic cell growth or proliferation. The products  
 CC can be used in the production of transgenic or knock-out animals.  
 CC Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing  
 CC cells.  
 XX S0 Sequence 1734 BP; 355 A; 491 C; 495 G; 393 T; 0 other;  
 Query Match 87.5%; Score 659; DB 20; Length 1734;  
 Best Local Similarity 93.0%; Pred. No. 7.4e-168;  
 Matches 702; Conservative 0; Mismatches 50; Indels 3; Gaps 1;  
 OY 1 ATGAGGGGACCCAGTATCATCTTGGCCACTTCTCTCTGCTTCTCTCAATG 60  
 DB 257 ATGAGGGGACCCAGTATCATCTTGGCCACTTCTCTCTGCTTCTCTCAATG 316  
 OY 61 GTGTGTGCGCCAGCTGTGCGGACACCTGTACTGTCTCTGACACACCCCATGCCCC 120  
 DB 317 GTGTATTCCTCAGCTGTGCGGACACCTGTCTCTCTGACACACCCCATGCCCC 376  
 OY 121 CAGGGGGTACCCCTGGTGTGATGGCTGTGGCTGTCTAAGTGTGACGGAGGCTG 180  
 DB 377 CCGGGGGTACCCCTGGTGTGATGGCTGTGGCTGTCTAAGTGTGACGGAGGCTG 436  
 OY 181 GGGGAGTCTCTGACACCATCTGATGTGACACCCAGCGGGCTGTGTCAGGCT 240  
 DB 437 GGGGAGTCTCTGACACCATCTGATGTGACACCCAGCGGGCTGTGTCAGGCT 496  
 OY 241 GGGGAGGCTCTGGGGCCATGGGGCTGTGTCTCTTGGATGAGGATGACGGTACTGT 300

DB 497 GGGGAGGCTCTGGGGCCAGTGTGCTGTGTGCTGTGCTCTTTCGAGAGGATGACGGAGCTGT 556  
 OY 301 GAGGTGATGAGCCCGCAGAGTACCTGATGATGAGAGACCTTTAAACCAATTGACGAGGCTCG 360  
 DB 557 GAGGTGATGAGCCCGCAGAGTACCTGATGATGAGAGACCTTTAAACCAATTGACGAGGCTCG 616  
 OY 361 TGCCGCTGTGATGACAGGTGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420  
 DB 617 TGCCGCTGTGATGACAGGTGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 676  
 OY 421 CCCAGCTGGGAGTGTGACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480  
 DB 677 CCCAGCTGGGAGTGTGACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 736  
 OY 481 TGCGTATGTGACACGAGGATGA---CACCGGCGATTCACAGCGCTCCACGCGCAGAGACAC 537  
 DB 737 TGCGTATGTGACACGAGGATGA---CACCGGCGATTCACAGCGCTCCACGCGCAGAGACAC 796  
 OY 538 CAACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 597  
 DB 797 CAACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 856  
 OY 598 GCGTGGGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 657  
 DB 857 GCGTGGGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 916  
 OY 658 AACCGATTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 717  
 DB 917 AACCGATTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 976  
 OY 718 GCCAGGAGCCACAGCTCATGGAACAGTCTTCTTA 752  
 DB 977 TCCAGGAGCCACAGCTCATGGAACAGTCTTCTTA 1011  
 RESULT 5  
 AAX76489/c  
 ID AAX76489 standard; DNA: 753 BP.  
 XX AAX76489;  
 DT 06-AUG-1999 (first entry)  
 XX DE Mouse WISP-2 protein complementary nucleotide sequence SEQ ID NO:18.  
 KW WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;  
 KW connective tissue growth factor; cancer; melanoma; arteriosclerosis;  
 KW leukemia; lymphoid malignancy; haematopiasis-related disorder;  
 KW tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;  
 KW kidney disorder; bone-related disorder; osteoporosis; trauma; burn;  
 KW connective tissue disorder; catabolic state; inflammation;  
 KW testicular-related disorder; angiogenesis; immunological disorder; ss.  
 XX OS Mus sp.  
 XX PN W09921998-A1.  
 XX PD 06-MAY-1999.  
 XX PE 29-OCT-1998; 98WO-US22991.  
 XX PR 14-APR-1998; 98US-0081695.  
 XX PR 29-OCT-1997; 97US-0063704.  
 XX PR 03-FEB-1998; 98US-0073612.  
 XX PA (GENETECH ) GENENTECH INC.  
 XX PI Botstein DA, Cohen RL, Goddard A, Gurney AL, Hillan K;  
 XX PI Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WI;  
 DR WPI: 1999-337420/28.

PT New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3  
 XX Disclosure: Page 179-180; 284pp; English.  
 XX  
 CC The present invention describes Wnt-1 induced secreted polypeptides,  
 CC WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2  
 CC and WISP-3 have homology to connective tissue growth factor (CTGF).  
 CC Products from the present invention can be used to treat WISP-related  
 CC disorders such as breast, ovarian, and colon cancer or melanoma. The  
 CC products can be used to treat arteriosclerosis. The products can also be  
 CC used to treat other diseases e.g. Dengn and malignant tumours,  
 CC leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,  
 CC hypothalamic and other glandular, macrophagal, epithelial, stromal, and  
 CC blastococelic disorders, haematopolesis-related disorders, tissue-growth  
 CC disorders, skin disorders, desmoplasia, fibrotic lesions, kidney  
 CC disorders, bone-related disorders such as osteoporosis, trauma such as  
 CC burns, incisions, and other wounds, connective tissue disorders,  
 CC catabolic states, testicular-related disorders, and inflammatory,  
 CC angiogenic and immunologic disorders including arteriosclerosis. The  
 CC products can also be used for detection and diagnosis especially of  
 CC individuals with neoplastic cell growth or proliferation. The products  
 CC can be used in the production of transgenic or knock-out animals.  
 CC Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing  
 CC cells.  
 XX  
 SQ Sequence 753 BP; 159 A; 229 C; 238 G; 127 T; 0 other;  
 Query Match 87.3%; Score 657; DB 20; Length 753;  
 Best Local Similarity 93.0%; Pred. No. 2e-167;  
 Matches 700; Conservative 0; Mismatches 50; Indels 3; Gaps 1;  
 QY 1 ATGAGGGGAGCCACCTGATTCATCTTGGCCACTTCCCTCTGCTCTCTCAANG 60  
 DB 753 ATGAGGGGAGCCACCTGATTCATCTTGGCCACTTCCCTCTGCTCTCTCAANG 60  
 QY 61 GTGTGTCCTGAGCTGTGCGGACACCTGTACCTGTCTTGGACACACCCAGTCCCA 120  
 DB 693 GTGTATTTCCAGCTGTGCGGACACCTGTGCTGTCTTGGACACACCCAGTCCCA 624  
 QY 121 CAGGGGGTACCCCTGGTGTGATGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 180  
 DB 633 CCGGGGGTACCCCTGGTGTGATGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 574  
 QY 181 GGGGAGTCCCTGGACACCTGATGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 240  
 DB 573 GGGGAGTCCCTGGACACCTGATGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 514  
 QY 241 GGGGAGGCTCTGGCGGACATGGGGCTGTGTCTCTTGGATGAGATGACGGTAGCTGT 300  
 DB 513 GGGGAGGCTCTGGCGGACATGGGGCTGTGTCTCTTGGATGAGATGACGGTAGCTGT 454  
 QY 301 GAGGTGAATGGCCGACAGTACCTGATGTGAGAGAGACTTTAAACCAATTGACAGGTTTG 360  
 DB 453 GAGGTGAATGGCCGACAGTACCTGATGTGAGAGAGACTTTAAACCAATTGACAGGTTTG 394  
 QY 361 TGCCCGCTGTGATGACAGTGGCTTACCTGCTCCCTGTGCTGTGAGATGATGGCGCTG 420  
 DB 393 TGCCCGCTGTGATGACAGTGGCTTACCTGCTCCCTGTGCTGTGAGATGATGGCGCTG 334  
 QY 421 CCCAGCTGTGAGTGGCCACGCCCCAAGAGATACAGGTGCCAGAGAAAGTCTGCCCGAG 480  
 DB 333 CCCAGCTGTGAGTGGCCACGCCCCAAGAGATACAGGTGCCAGAGAAAGTCTGCCCGAG 274  
 QY 481 TGGGTATGTGACAGGAGTGA---CAGCGGCGATCCAGCGCTCCAGCGCGCAGAGACAC 537  
 DB 273 TGGGTATGTGACAGGAGTGA---CAGCGGCGATCCAGCGCTCCAGCGCGCAGAGACAC 214  
 QY 538 CAACCTTCTGCCCTTCTCACTCCTGCTGCTGATGATGCTCTGTGCTCAATTGGAGACA 597  
 DB 213 CAACCTTCTGCCCTTCTCACTCCTGCTGATGATGCTGCGGATGCGCCCTGTCAAACTGGAGACA 154  
 QY 598 GCGTGGGGCCCTGTCAACACACCTGTGGGCTGGGATAGCCACCGGAGTGTCAACACAG 657  
 |||||||

DB 153 GCGTGGGGCCCTGTCAACACACCTGTGGGTGGGATAGCCACCGGATATCAACACAG 94  
 QY 658 AACGATTCGCACTGAGATGATCAACGCGCGCTGTGCTGCGCAGACCCCTCCGGA 717  
 DB 93 AACGATTCGCACTGAGATGATCAACGCGCGCTGTGCTGCGCAGACCCCTCCGGA 34  
 QY 718 GCCAGAGCCACAGCTCATGATGAGACAGTGTTC 750  
 DB 33 TCCAGAGCCACAGCTCATGATGAGACAGTGTTC 1  
 RESULT 6  
 ID AAX28435  
 XX AAX28435 standard; DNA; 1257 BP.  
 AC AAX28435;  
 XX  
 XX 22-JUN-1999 (first entry)  
 DE EGF-like homologue PRO261 coding sequence.  
 XX  
 KW Antibody; PRO187; PRO533; PRO214; PRO240; PRO211; PRO230; PRO261; PRO246;  
 KM EBAF-2; Inhibitor; tumour growth; cancer; EGF-like homologue;  
 KM EGF-8 homologue; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9914327-A2.  
 XX  
 PD 25-MAR-1999.  
 XX  
 PF 10-SEP-1998; 98MO-US18824.  
 XX  
 PR 25-NOV-1997; 97US-0066840.  
 PR 17-SEP-1997; 97US-0059114.  
 PR 17-SEP-1997; 97US-0059117.  
 PR 18-SEP-1997; 97US-0059263.  
 PR 15-OCT-1997; 97US-0062125.  
 PR 17-OCT-1997; 97US-0062285.  
 PR 17-OCT-1997; 97US-0062287.  
 PR 24-OCT-1997; 97US-0062816.  
 PR 29-OCT-1997; 97US-0063704.  
 XX  
 PA (GETH ) GENENTECH INC.  
 XX  
 PI Botstein D, Goddard A, Gurney A, Hillan K, Lawrence DA;  
 PI Roy M, Wood WI;  
 XX  
 DR WPI: 1999-229532/19.  
 DR P-PSDB; AAY05285.  
 XX  
 PT Antibodies against specific proteins overexpressed in tumours  
 XX  
 PS Example 1; Fig 22; 130pp; English.  
 XX  
 CC This sequence encodes the EGF-like homologue PRO261.  
 CC The invention relates to antibodies (Ab) that bind to any of the  
 CC polypeptides (I) designated PRO187, PRO533; PRO214; PRO240; PRO211;  
 CC PRO230; PRO261; PRO246 or EBAF-2. The Ab, or other agents that inhibit  
 CC expression and/or activity of (I) are used: (i) to inhibit growth of  
 CC tumours; and (ii) as diagnostic/prognostic reagents for detection or  
 CC quantification of (I) in cells or tissues, by standard immunoassays, with  
 CC overexpression being indicative of cancer. For therapeutic use, the Ab  
 CC may be conjugated to a toxin, chemotherapeutic agent or radioisotope.  
 CC Genes expressing (I), many of which are growth factor homologues, are  
 CC overexpressed in some cases of cancer.  
 XX  
 SQ Sequence 1257 BP; 215 A; 416 C; 385 G; 241 T; 0 other;  
 Query Match 67.8%; Score 510.4; DB 20; Length 1257;  
 Best Local Similarity 79.9%; Pred. No. 8.2e-128;  
 Matches 601; Conservative 0; Mismatches 151; Indels 0; Gaps 0;





```

OY 481 TGGGTATGTGACAGGAGTGTACACCGGCGATGACGGCTCCACGCGCAAGACACCA 540
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 490 TGGGTGTGCGGCAAGAGGGGACTGGGGACCGACCGCTTCCAGCCCAAGACCCAG 549
OY 541 CTTTGTGCGCTTGTACTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 550 TTTTCTGCGCTTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 609
OY 601 TGGGGCCCTGCTCAACCACTGTGGCTGGGCTAGGACACCGAGTGTCCAAACCAAC 660
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 610 TGGGACCTCTGTCCACCACTGTGGCTGGGCTAGGACACCGAGTGTCCAAACCAAC 669
OY 661 CGATTCTGCCAATGTGAGATTCACACCGCGCTGTGCTGCTGCTGCTGCTGCTGCTG 720
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 670 CGCTTCTGCGAGTGTGAGACCGACCGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 729
OY 721 AGGACCCACAGCTCATGGAACAGTCTTTCTA 752
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 730 AGGGGTGCGAGTGTCCACAAACAGTCTTTCTA 761

```

## RESULT 9

AAE60368 ID AAE60368 standard; cDNA; 1266 BP.

AAE60368;

27-APR-2001 (first entry)

PRO261 coding sequence.

CYTOSTATIC: PRO protein; tumour; cancer; ss.

Homo sapiens.

W020105836-A1.

25-JAN-2001.

20-DEC-1999; 99MO-US30999.

20-JUL-1999; 99US-0144758.

26-JUL-1999; 99US-0145698.

08-SEP-1999; 99MO-US20594.

13-SEP-1999; 99MO-US20944.

15-SEP-1999; 99MO-US21090.

05-OCT-1999; 99MO-US23089.

29-NOV-1999; 99MO-US28214.

30-NOV-1999; 99MO-US28313.

02-DEC-1999; 99MO-US28564.

(GETH) GENENTECH INC.

Botstein D, Goddard A, Gurney AL, Hillan KJ, Roy MA, Wood WI;

WPI, 2001-091968/10.

P-PSDB; AAB68598.

New antibody that binds to a PRO polypeptide, e.g. PRO187 and PRO533,

useful for diagnosing and treating cancers -

Claim 50; Fig 13; 196pp; English.

The present invention relates to PRO proteins and coding sequences. The

present sequence is the coding sequence for one such PRO protein.

It was found that the PRO genes are amplified in the genome of tumour

cells. The gene amplification is expected to be associated with the

overexpression of the gene product and contributes to tumourigenesis.

Therefore, antagonists of PRO proteins are useful for the treatment of

benign or malignant tumours, leukemias, lymphoid malignancies and other

disorders such as neuronal, glial, astrocytic, hypochalamic, glandular,

epithelial, inflammatory and immunologic disorders.

XX SQ Sequence 1266 BP; 216 A; 418 C; 390 G; 242 T; 0 other;

Query Match 67.8%; Score 510.4; DB 22; Length 1266;

Best Local Similarity 79.9%; Pred. No. 8.3e-128;

Matches 601; Conservative 0; Mismatches 151; Indels 0; Gaps 0;

```

OY 1 ATGAGGGGACGGCCATGATTCATCTTGTGGCACTTCTCTGCTGCTCTCTCAATG 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 10 ATGAGGAGACACGAGACCCACCTCTGGCTTCTCTGCTGCTCTCTCAATG 69
OY 61 GTGTGTGCGGCTGTGCGGACACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 70 GTGCTTACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 129
OY 121 CAGGGGATACCCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 130 CTGGAGATACCCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 189
OY 181 GGGGAGTCTGCGACACCTGCATGTGCGACCCACGACGAGGCTGCTGCTGCTGCTGCT 240
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 190 GGGGAGCCCTGCGACCACTGCATGTGCGACGAGGCTGCTGCTGCTGCTGCTGCTGCTGCT 249
OY 241 GGGGAGGCGCTGCGGCGCATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 250 GGGGAGGACCCGCTGCGGCGCATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 309
OY 301 GAGTGAATGCGCGGACCTGATGAGAGAGACCTTAAACCAATTGACAGGCTCTG 360
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 310 GAGGTGAACGCGCGCTGATGAGAGAGAGACCTTAAACCAATTGACAGGCTCTG 369
OY 361 TGCCGCTGTGATGACGCTGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 370 TGCCGCTGTGATGACGCTGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 429
OY 421 CCCAGCTGGAGTGTGCGGCGCATGAGAGAGAGACCTTAAACCAATTGACAGGCTCTG 480
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 430 CCCAGCTGGAGTGTGCGGCGCATGAGAGAGAGACCTTAAACCAATTGACAGGCTCTG 489
OY 481 TGGGTATGTGACAGGAGTGTACACCGGCGATGACCGCTCCACGCGCAAGACACCA 540
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 490 TGGGTGTGCGGCAAGAGGGGACTGGGGACCGACCGCTTCCAGCCCAAGACCCAG 549
OY 541 CTTTGTGCGCTTGTACTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 550 TTTTCTGCGCTTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 609
OY 601 TGGGGCCCTGCTCAACCACTGTGGCTGGGCTAGGACACCGAGTGTCCAAACCAAC 660
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 610 TGGGACCTCTGTCCACCACTGTGGCTGGGCTAGGACACCGAGTGTCCAAACCAAC 669
OY 661 CGATTCTGCCAATGTGAGATTCACACCGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 720
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 670 CGCTTCTGCGAGTGTGAGACCGACCGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 729
OY 721 AGGACCCACAGCTCATGGAACAGTCTTTCTA 752
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 730 AGGGGTGCGAGTGTCCACAAACAGTCTTTCTA 761

```

## RESULT 10

AAC97451 ID AAC97451 standard; cDNA; 1266 BP.

AAC97451;

28-FEB-2001 (first entry)

Human angiogenesis-associated protein PRO261 cDNA, SEQ ID NO:107.

Human; angiogenesis-associated protein; PRO; endothelial cell growth;

cardiac hypertrophy; cardiovascular disorder; endothelial disorder;

angiogenic disorder; atherosclerosis; osteoporosis; hypertension;



KM Connective tissue growth factor-3; CTGF-3; human; cancer;  
 KM arthritis; fibrosis; osteoporosis; diagnosis; therapy; ds.  
 OS Homo sapiens.  
 XX Key Location/Qualifiers  
 FH CDS 9..761  
 FT /\*tag= a  
 FT s1g\_peptide 9..65  
 FT /\*tag= b  
 FT mat\_peptide 66..758  
 FT /\*tag= c  
 XX W09621236-A1.  
 XX 22-MAY-1998.  
 XX 08-NOV-1996; 96WO-US17856.  
 XX 08-NOV-1996; 96WO-US17856.  
 XX 08-NOV-1996; 96WO-US17856.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX Chopra A, Ebner R, Ruben SM;  
 XX WPI: 1998-297864/26.  
 XX P-PSDB; AAW37946.  
 PT Novel human connective tissue growth factor 3 gene - useful for the  
 PT diagnosis and treatment of e.g. cancer, arthritis, fibrosis,  
 PT osteoporosis  
 XX  
 PS Claim 2; Fig 1A-B; 87bp; English.  
 CC This nucleotide sequence codes for human connective tissue growth  
 CC factor-3 (CTGF-3) protein (see AAW37946), a novel member of the  
 CC growth factor superfamily. It was discovered in a cDNA library  
 CC derived from human osteoblasts. The gene has also been identified  
 CC in cDNA libraries from ovary, testis, heart, lung, skeletal muscle,  
 CC adrenal medulla, adrenal cortex, thymus, prostate, small intestine  
 CC and colon. A cDNA clone is deposited as ATCC 97756. Also provided  
 CC are vectors, host cells and recombinant methods for producing  
 CC CTGF-3 polypeptides. CTGF-3 nucleic acid sequences or their  
 CC fragments, e.g. primers or probes, can be used to diagnose diseases  
 CC where CTGF-3 expression is enhanced, e.g. cancer, arthritis,  
 CC fibrosis or atherosclerosis, or diseases where expression is  
 CC decreased such as in osteoporosis. Disorders characterised by  
 CC decreased or increased levels of CTGF-3 can be treated by  
 CC administering CTGF-3 polypeptides and anti-CTGF-3 antibodies,  
 CC respectively.  
 CC  
 CC Sequence 1285 BP; 237 A; 418 C; 389 G; 241 T; 0 other;  
 SO  
 Query Match 67.8%; Score 510.4; DB 19; Length 1285;  
 Best Local Similarity 79.9%; Pred. No. 8.3e-128;  
 Matches 601; Conservative 0; Mismatches 151; Indels 0; Gaps 0;

QY 241 GGGGAGGAGCCCTGGCGCCATGAGGGGCTGTGTCTCTTGATGAGATGAGCTAGCTGT 300  
 DB 249 GGGGAGGAGCCCGGCGGCGCCGCGGCGCTGTGCTCTTGAGAGAGAGACAGAGCTGT 308  
 QY 301 GAGGTGATGAGCCGAGTACCTGATGAGAGACCTTAAACCAATTGACGGCTCTG 360  
 DB 309 GAGGTGATGAGCCGCTGTATGAGGAGGAGGAGCTTCAAGCCCATGTCAGACATCCGC 368  
 QY 361 TGCCGCTGTGATGAGGAGGAGGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTG 420  
 DB 369 TGCCGCTGTGAGAGAGGAGGAGGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTG 428  
 QY 421 CCCAGCTGTGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 480  
 DB 429 CCCAGCTGTGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 488  
 QY 481 TGGGTATGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540  
 DB 489 TGGGTATGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 548  
 QY 541 CTTTCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
 DB 549 TTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 608  
 QY 601 TGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660  
 DB 609 TGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 668  
 QY 661 CGATTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720  
 DB 669 CGATTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 728  
 QY 721 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 752  
 DB 729 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 760  
 RESULT 12  
 AAX76486  
 ID AAX76486 standard; DNA; 1293 BP.  
 AC AAX76486;  
 XX  
 DT 06-AUG-1999 (first entry)  
 XX  
 DE Human WISP-2 protein nucleotide sequence SEQ ID NO:13.  
 XX  
 KW WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;  
 KW connective tissue growth factor; cancer; melanoma; arteriosclerosis;  
 KW leukaemia; lymphoid malignancy; hematopoiesis-related disorder;  
 KW tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;  
 KW kidney disorder; bone-related disorder; osteoporosis; trauma; burn;  
 KW connective tissue disorder; catabolic state; inflammation;  
 KW testicular-related disorder; angiogenesis; immunological disorder; ss.  
 XX  
 OS Homo sapiens.  
 XX W09921998-A1.  
 XX 06-MAY-1999.  
 XX 29-OCT-1998; 96WO-US22991.  
 XX 14-APR-1998; 98US-0081695.  
 XX 29-OCT-1997; 97US-0063704.  
 XX 03-FEB-1998; 98US-0073612.  
 XX (GENE) GENENTECH INC.  
 XX Botstein DA, Cohen RL, Goddard A, Gurney AL, Hillan K;  
 XX Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WT;  
 XX WPI: 1999-337420/28.



DR P-PSDB: AA17649.

PT New Isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3

XX Example 4; Page 174-175; 284pp; English.

XX The present invention describes Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2 and WISP-3 have homology to connective tissue growth factor (CTGF). CC Products from the present invention can be used to treat WISP-related disorders such as breast, ovarian, and colon cancer or melanoma. The CC products can be used to treat arteriosclerosis. The products can also be used to treat other diseases e.g. benign and malignant tumors, leukaemia and lymphoid malignancies, neuronal, glial, astrocytic, and hypothalamic and other glandular, macrophage, epithelial, stromal, and CC blastocoele disorders, hematopoiesis-related disorders, tissue growth disorders, skin disorders, desmoplasia, fibrotic lesions, kidney CC disorders, bone-related disorders such as osteoporosis, trauma such as burns, incisions, and other wounds, connective tissue disorders, CC catabolic states, testicular-related disorders, and inflammatory, CC angiogenic and immunologic disorders including arteriosclerosis. The CC products can also be used for detection and diagnosis especially of CC individuals with neoplastic cell growth or proliferation. The products can be used in the production of transgenic or knock-out animals. CC Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing CC cells.

XX Sequence 1293 BP; 232 A; 425 C; 393 G; 243 T; 0 other;

Query Match 67.8%; Score 510.4; DB 20; Length 1293;

Best Local Similarity 79.9%; Pred. No. 8.3e-128; Matches 601; Conservative 0; Mismatches 151; Indels 0; Gaps 0;

QY 1 ATGGGGGAGCCAGCCATGATCCATCTGCGCCACTCTCTCTCTCTCTCTCTCATG 60  
DB 22 ATGGAGAGGACACCGAAGACCCACTCTCTCTCTCTCTCTCTCTCTCTCTCAAG 81  
QY 61 GTGTGTGCCAGCTGTGTGCGGACACCTGTACTCTCTCTGAGACACCCAGTGGCCA 120  
DB 82 GTGGGTACCCAGCTGTGTGCGGACACCATGTACTCTCTCTGCGCCACTCTCTCTG 141  
QY 121 CAGGGGTACCCCTGTGTGTGATGCTGTGTGCTGTAAAGTGTGACGAGGCTG 180  
DB 142 CTGGGAGTACCCCTGTGTGTGATGCTGTGTGCTGTGCGGATGTGACAGCGGCTG 201  
QY 181 GGGGAGTCCGCGACACCATGCTGTGACCCCGACGAGGCTCTGTGTGACCTT 240  
DB 202 GGGGAGCTGTGCGACACTGTGCTGTGACGCGACGAGGCTGTGTGACGAGCC 261  
QY 241 GGGGAGGCTCTGTGCGACATGAGGCTGTGTGTGATGAGATGAGGAGTGT 300  
DB 262 GGGGAGGACCCGCTGTGCGGAGGCTGTGTGTGATGAGATGAGGAGTGT 321  
QY 301 GAGGTAAATGCGCGACGATCTGTGATGAGAGACCTTTAAACCAATGCGAGGCTCTG 360  
DB 322 GAGGTAAATGCGCGACGATCTGTGATGAGAGACCTTTAAACCAATGCGAGGCTCTG 381  
QY 361 TGCGGTGTGATGAGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 420  
DB 382 TGCGGTGTGATGAGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 441  
QY 421 CCCAGCTGTGAGTCCCGACGCGCCAGAGATACAGTGTGCGAGAAATGCGCGGAG 480  
DB 442 CCCAGCTGTGAGTCCCGACGCGCCAGAGATACAGTGTGCGAGAAATGCGCGGAG 501  
QY 481 TGGGTATGTACCAAGGAGTGTGACACCGCGATCCAGCGCTCCAGCGACAGACCA 540  
DB 502 TGGGTATGTACCAAGGAGTGTGACACCGCGATCCAGCGCTCCAGCGACAGACCA 561  
QY 541 CT 600  
DB 562 TTTTCTGTGCTGTCT 621

QY 601 TGGGCCCCCTGTCTCAACCACTGTGTGCGCTGGGATATACCCGAGTGTCCACAGAAC 660  
DB 622 TGGGAGACCTGTCTGTACACCTGTGTGCGCTGGGATATACCCGAGTGTCTCCACAGAAC 681  
QY 661 CGATTCTGCACTGTGAGATTCACAGCGGCTGTGTGCTGCCACAGACCTCTGGACGCC 720  
DB 682 CGCTTCTGCGACGTGAGACCCAGCGGCTGTGTGCTGTGCTGCCACAGACCTCTGGACGCC 741  
QY 721 AGGAGCCACAGCTCATATGAGACAGTCTTCTTA 752  
DB 742 AGGGGTGCGAGTCCACCAAGAGTCTCTTA 773

#### RESULT 13

ID AAH28214 standard; cDNA: 1309 BP.

XX AAH28214;

DT 05-SEP-2001 (first entry)

DE Connective tissue derived growth factor related protein cDNA.

XX Growth factor; protein inhibitor; protease; damaged tissue;  
KW platelet-derived growth factor; PDGF; fibroblast growth factor; FGF;  
KW connective tissue derived growth factor; CTGF; chrysalin; VEGF;  
KW keratinocyte-derived growth factor; KGF; epidermal growth factor; EGF;  
KW transforming growth factor-beta; TGF-beta; matrix metalloproteinase; MMP;  
KW granulocyte macrophage colony stimulating factor; GM-CSF; uPA;  
KW vascular endothelial growth factor; urokinase plasminogen activator;  
KW dermal ulcer; wound; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 7..759

XX FT /\*tag= a /product= "connective tissue derived growth factor related protein"

XX W0200149309-A2.

XX 12-JUL-2001.

XX 21-DEC-2000; 2000WO-IB01935.

XX 29-DEC-1999; 99GB-0030768.

XX (PFIZ ) PFIZER LTD.

XX (PFIZ ) PFIZER INC.

XX Davies MJ, Huggins JP, McIntosh FS, Ocleston NL;

XX WPI: 2001-418351/44.

XX P-PSDB: AAB84599.

XX Composition for the treatment of damaged tissue i.e. chronic wounds and  
PT dermal ulcers comprises an inhibitor agent i.e. a protease and a growth  
PT factor -

XX Disclosure: Page 546; 572pp; English.

XX The specification describes a pharmaceutical composition, comprising  
CC a growth factor, an inhibitor agent, i.e. a protease. The inhibitor  
CC agent inhibits the action of at least one specific adverse protein,  
CC i.e. a protease, that is upregulated in a damaged tissue such as a  
CC wound environment. Growth factors which are included in the composition  
CC of the invention are platelet-derived growth factor (PDGF), fibroblast  
CC growth factor (FGF), connective tissue derived growth factor (CTGF),  
CC keratinocyte-derived growth factor (KGF), transforming growth  
CC factor-beta (TGF-beta), granulocyte macrophage colony stimulating factor  
CC (GM-CSF), epidermal growth factor (EGF), vascular endothelial growth  
CC factor (VEGF), and chrysalin. Inhibitors which are included in the

CC composition of the invention include inhibitors of urokinase-type  
 CC plasminogen activator (uPA) and matrix metalloproteinase (MMP). The  
 CC composition is useful for the treatment of chronic damaged tissue, i.e.  
 CC wounds and dermal ulcers. The present sequence encodes a human  
 CC CTGF-like protein, and is used to produce the composition of the  
 CC invention.

XX Sequence 1309 BP; 261 A; 418 C; 387 G; 242 T; 1 other:

Query Match 67.8%; Score 510.4; DB 22; Length 1309;  
 Best Local Similarity 79.9%; Pred. No. 8.4e-128;  
 Matches 601; Conservative 0; Mismatches 151; Indels 0; Gaps 0;

```

OY 1 ATGAGGGGACCCCACTGATCATCTTCTGGCCACTTCTCTGCTTCTCTCAATG 60
DB 7 ATGAGAGGACACCGAAGACCCACTCTGCGCTTCTCTGCTCTCTCTCTCAAG 66
OY 61 GTGTGTGCCAGCTGTGCGGACACCTGTACCTGTCTTGACACCAACCCCAAGTCCA 120
DB 67 GTGGGTACCCAGCTGTGCGGACACCAATGTACCTGCGCTGCGGACCTGATGCCG 126
OY 121 CAGGGGTACCCCTGTGATGCTGTGCTGTGCTGTAAAGTGTGACGAGGAGCTG 180
DB 127 CTGGGAGTACCCCTGTGCTGTGATGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 186
OY 181 GGGAGTCTCGACACCACTGATGTGCGACCCGACGAGGCGCTGTTGTACAGCT 240
DB 187 GGGAGGCGCTGACCACTGACCTGACGTGTGCGACGCGACGAGGCGCTGTGCGACG 246
OY 241 GGGGAGGCGCTGCGGCGGACGAGGCGCTGTGTCTTGTGATGAGGATGACGTAAGT 300
DB 247 GGGGAGGAGCGCGGCGGCGGCGGCGCTGTGCTGTGCGACGAGGAGGAGCGTGT 306
OY 301 GAGGTGAATGGCGCGGACCTGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360
DB 307 GAGGTGAATGGCGCGCTGTATGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 366
OY 361 TGCCGCTGTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 420
DB 367 TGCCGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 426
OY 421 CCCAGCTGGAGCTGCGGCGGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 480
DB 427 CCCAGCTGGAGCTGCGGCGGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 486
OY 481 TGGGATGTGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540
DB 487 TGGGATGTGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 546
OY 541 CTTTCTGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 600
DB 547 TTTTCTGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 606
OY 601 TGGGGCGGCTGTCAACCACTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660
DB 607 TGGGGCGGCTGTCAACCACTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 666
OY 661 CGATTCTGCAACTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 720
DB 667 CGATTCTGCAACTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 726
OY 721 AGGAGCCACAGCTCATGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 752
DB 727 AGGAGCCACAGCTCATGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 758

```

## RESULT 14

AA1595 standard; cDNA: 1522 BP.

XX AA1595;  
 XX AA1595;  
 DT 29-APR-1999 (first entry)

XX Human growth factor-like protein encoding cDNA.  
 DE Human: growth factor-like protein; HGFLP: fisp-12; CTGF: PDGF family;  
 XX connective tissue growth factor; platelet-derived growth factor; keloid;  
 KW connective tissue disorder; cancer; ankylosing spondylitis; scleroderma;  
 KW atherosclerosis; Dupuytren's contracture; eosinophilic fasciitis;  
 KW Fely syndrome; Goodpasture's disease; Hunter syndrome; Hurler syndrome;  
 KW Marfan syndrome; nodular fasciitis; osteogenesis imperfecta; restenosis;  
 KW rheumatoid arthritis; systemic lupus erythematosus; ss.  
 XX  
 XX Homo sapiens.  
 OS  
 XX  
 XX Key Location/Qualifiers  
 FH 256..1008  
 FT CDS /tag= a  
 FT

```

XX W09858063-A1.
XX 23-DEC-1998.
XX 18-JUN-1998; 98WO-US12787.
XX 19-JUN-1997; 97US-0878990.
XX (INCY-) INCYTE PHARM INC.
XX Corley NC, Hillman JL, Lal P;
XX WPI: 1999-080954/07.
XX P-PSDB: AAM94616.
XX New polynucleotide encoding growth factor-like protein - useful for
XX preventing and treating connective tissue disorders and cancer
XX Claim 4; Fig 1; 63pp; English.
XX
XX The present sequence encodes human growth factor-like protein (HGFLP)
XX which has homology to connective tissue growth factor (CTGF) and fisp-12,
XX which are members of the platelet-derived growth factor (PDGF)
XX superfamily of growth factors. Antisense oligonucleotides and HGFLP
XX antagonists can be used to treat or prevent cancer, and connective
XX tissue disorders including ankylosing spondylitis, atherosclerosis,
XX Dupuytren's contracture, eosinophilic fasciitis, Fely syndrome,
XX Goodpasture's disease, Hunter syndrome, Hurler syndrome, Marfan
XX syndrome, nodular fasciitis, osteogenesis imperfecta, polyarthrits
XX nodosa, rheumatoid arthritis, scleroderma, systemic lupus erythematosus,
XX and restenosis following angioplasty. HGFLP antibodies and
XX polynucleotides can also be used in diagnostic assays for conditions or
XX diseases characterised by GRFLP expression.
XX
XX Sequence 1522 BP; 279 A; 496 C; 459 G; 288 T; 0 other:

```

Query Match 67.8%; Score 510.4; DB 20; Length 1522;  
 Best Local Similarity 79.9%; Pred. No. 8.8e-128;  
 Matches 601; Conservative 0; Mismatches 151; Indels 0; Gaps 0;

```

OY 1 ATGAGGGGACCCCACTGATCATCTTCTGGCCACTTCTCTGCTTCTCTCAATG 60
DB 256 ATGAGAGGACACCGAAGACCCACTCTGCGCTTCTCTGCTCTCTCTCTCAAG 315
OY 61 GTGTGTGCCAGCTGTGCGGACACCTGTACCTGTCTTGACACCAACCCCAAGTCCA 120
DB 316 GTGGGTACCCAGCTGTGCGGACACCAATGTACCTGCGCTGCGGACCTGATGCCG 375
OY 121 CAGGGGTACCCCTGTGATGCTGTGCTGTGCTGTAAAGTGTGACGAGGAGCTG 180
DB 376 CTGGGAGTACCCCTGTGATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 435
OY 181 GGGGAGTCTCGACACCACTGATGATGCGACCCGACGAGGCGCTGTTGTACAGCT 240
DB 436 GGGGAGGCGCTGCGACCACTGACGTGTGCGACGCGACGAGGCGCTGTGCTGCGACG 495

```

QY 241 GGGGAGAGCCCTGGGCGCATGGGGCTGTGTCTCTTGTGATGAGATGACGTAGCTGT 300  
 DB 496 GGGGAGAGCCCGGTGGGCGGGGCGCTGTGCTTGTGGAGAGACGACAGAGCTGT 555  
 QY 301 GAGGTGAATGGCCGAGTACTGTGATGAGAGACCTTTAAACCAATTGAGGCTGT 360  
 DB 556 GAGGTGAAGCGCCCTGTATCCGGAAAGGGAGACCTTCCAGCCCACTGACATCCGC 615  
 QY 361 TGCCGCTGTGATGAGGCTGGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420  
 DB 616 TGCCGCTGTGAGAGAGCGGCGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 675  
 QY 421 CCCAGCTGGAGACTGCCAGCGCCCAAGAGATACAGGTGCCAGAAAGTGTGCCCGAG 480  
 DB 676 CCCAGCTGGAGACTGCCAGCGCCCAAGAGATACAGGTGCCAGAAAGTGTGCCCGAG 735  
 QY 481 TGGGTATGTACACGAGGAGTACACCGCGCATCCAGGCTCCAGCGCCAGACCA 540  
 DB 736 TGGGTATGTACACGAGGAGTACACCGCGCATCCAGGCTCCAGCGCCAGACCA 595  
 QY 541 CTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600  
 DB 796 TTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 855  
 QY 601 TGGGGCCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660  
 DB 856 TGGGGCCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 915  
 QY 661 CGATTCTGCGCACTGAGATGCCAGCGCGCTGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 720  
 DB 916 CGATTCTGCGCACTGAGATGCCAGCGCGCTGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 975  
 QY 721 AGGAGCCACATCATGTAACAGTCTTTCTA 752  
 DB 976 AGGAGCCACATCATGTAACAGTCTTTCTA 1007

RESULT 15  
 AAH46952  
 ID AAH46952 standard; cDNA: 1337 BP.

AC AAH46952;

DT 25-SEP-2001 (first entry)

DE Human secreted protein encoding cDNA (clone id HB0DE48).

KW Secreted protein; immunosuppressive; antiarthritic; antirheumatic;

KW antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;

KW neurotrophic; neuroprotective; antibacterial; virucide; fungicide; human;

OS opthalmological; gene therapy: ss.

XX Homo sapiens.

XX WO20015430-A1.

XX 02-AUG-2001.

XX 17-JAN-2001: 2001WO-US01431.

XX 31-JAN-2000: 2000US-0179065.

XX 04-FEB-2000: 2000US-0180628.

XX 12-SEP-2000: 2000US-0231968.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Komatsoulis GA, Baker KP, Birse CE, Soppet DR, Olsen HS;

XX Moore PA, Wei P, Ebner R, Duan DR, Shi Y, Choi GH, Fiscella M;

XX NI J, Ruben SM, Barash SC;

XX WPI: 2001-476220/51.

XX P-PSDB: AAB85542.

PT 17 isolated nucleic acid molecules encoding human secreted proteins,  
 used to preventing, treating or ameliorating a medical condition  
 XX  
 PS Claim 1: Page 427-428; 482pp; English.  
 CC The invention provides novel human secreted proteins and polynucleotides  
 CC encoding them. The secreted proteins can be expressed by standard  
 CC recombinant methodology. The secreted proteins and polynucleotides are  
 CC used to prevent, treat or ameliorate a medical condition in e.g. humans,  
 CC mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They can  
 CC also be used in diagnosing a pathological condition. The antibodies to  
 CC the proteins can also be used in alleviating symptoms associated with the  
 CC disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme  
 CC linked immunosorbent assays (ELISA). Disorders which are diagnosed or  
 CC treated include autoimmune diseases e.g. rheumatoid arthritis,  
 CC hyperproliferative disorders e.g. neoplasms of the breast or liver,  
 CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders  
 CC e.g. cerebral ischemia, angioneuromatosis, nervous system disorders e.g.  
 CC Alzheimer's disease, infections caused by bacteria, viruses and fungi and  
 CC ocular disorders e.g. corneal infection. The polypeptides can also be  
 CC used to aid wound healing and epithelial cell proliferation, to prevent  
 CC skin aging due to sunburn, to maintain organs before transplantation, for  
 CC supporting cell culture of primary tissues, to regenerate tissues and in  
 CC chemotaxis. The polypeptides can also be used as a food additive or  
 CC preservative to increase or decrease storage capabilities. The present  
 CC sequence represents a human secreted protein encoding cDNA.  
 XX  
 SO Sequence 1337 BP; 257 A; 427 C; 396 G; 252 T; 5 other:

Query Match 67.7%; Score 510; DB 22; Length 1337;  
 Best Local Similarity 79.8%; Pred. No. 1.1e-127;  
 Matches 600; Conservative 1; Mismatches 151; Indels 0; Gaps 0;

QY 1 ATGAGGGGAGCCGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 60  
 DB 47 ATGAGGGGAGCCGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 106  
 QY 61 GTGTGCGCCAGCTGTGCGCCGACACCCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120  
 DB 107 GTGTGCGCCAGCTGTGCGCCGACACCCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 166  
 QY 121 CAGGGGGTACCCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180  
 DB 167 CTGGGAGATACCCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 226  
 QY 181 GGGAGTCCCTGACACCCCTGATGCTGACACCCGACCCGACCCGCTTGTGACGCT 240  
 DB 227 GGGAGTCCCTGACACCCCTGATGCTGACACCCGACCCGACCCGCTTGTGACGCT 286  
 QY 241 GGGGAGCCCTGACACCCCTGATGCTGACACCCGACCCGACCCGCTTGTGACGCT 300  
 DB 287 GGGGAGCCCTGACACCCCTGATGCTGACACCCGACCCGACCCGCTTGTGACGCT 346  
 QY 301 GAGGTGAATGCGCCGAGTACCTGTGATGAGAGACCTTTAAACCAATTGAGGCTCTG 360  
 DB 347 GAGGTGAATGCGCCGAGTACCTGTGATGAGAGACCTTTAAACCAATTGAGGCTCTG 406  
 QY 361 TGCCGCTGTGATGAGAGGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420  
 DB 407 TGCCGCTGTGATGAGAGGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 466  
 QY 421 CCCAGCTGGAGACTGCCAGCGCCCAAGAGATACAGGTGCCAGAAAGTGTGCCCGAG 480  
 DB 467 CCCAGCTGGAGACTGCCAGCGCCCAAGAGATACAGGTGCCAGAAAGTGTGCCCGAG 526  
 QY 481 TGGGTATGTACACGAGGAGTACACCGCGCATCCAGGCTCCAGCGCCAGACCA 540  
 DB 527 TGGGTATGTACACGAGGAGTACACCGCGCATCCAGGCTCCAGCGCCAGACCA 586  
 QY 541 CTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600  
 DB 587 TTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 646

Fri Aug 1 10:44:21 2003

us-10-010-408-3.rng

Page 14

| QY | 601 | TGGGGCCCTGCTCAACCAACCGTGTGGGTGGGCATAGCAACCGGAGTGTCCAAACCGAAGC             | 660 |
|----|-----|---|-----|
|    |     |   |     |
| Db | 647 | TGGGGACCTCTCCACCAACCGTGTGGGTGGGCATAGCAACCGGAGTGTCCAAACCGAAGC <th>706</th> | 706 |
|    |     |   |     |
| QY | 661 | CGATTCTGCACATGGAGATCCAAAGCCGCGCTGTGCTGCAGAACCTGCCTGGAGAGC <th>720</th>    | 720 |
|    |     |   |     |
| Db | 707 | CGCTTCTCCGAGCTGGAGACCCAGGCGCGCTGTGCTGTCCAGAGCCCTGCACACCTCC <th>766</th>   | 766 |
|    |     |   |     |
| QY | 721 | AGGAGCCACAGCTCATGGACAGTCTTTCTA <th>752</th>                               | 752 |
|    |     |   |     |
| Db | 767 | AGGGTGCCAGTCCACAAACAGTCTTTCTA <th>798</th>                                | 798 |

Search completed: July 23, 2003, 11:15:20  
Job time : 200.398 secs



```

Db 437 GGGAGTCTCGACACCTGATGTCTGCCAGCCCAAGCCAGGGCCTGTTTACGCT 496
Qy 241 GGGGAGGCGCTGGGCGCATGGGCTGTGTCTTGGATGAGGATGAGGTACTGT 300
Db 497 GGGGAGGCGCGGAGGCTGTGTCTGTGTCTTGGAGAGGATGAGGGAGGTCT 556
Qy 301 GAGGTGATGGCCGAGGTACTGTGATGGAGAGACCTTAAACCAATTGACAGGTCTG 360
Db 557 GAGGTGATGGCCGAGGTACTGTGATGGAGAGACCTTAAACCAATTGACAGGTCTG 616
Qy 361 TGCCGCTGTGATGAGGTGTGCTTCACTGTGCGCTGTGCTGAGTGTGAGGTG 420
Db 617 TGCCGCTGTGATGAGGTGTGCTTCACTGTGCGCTGTGCTGAGTGTGAGGTG 676
Qy 421 CCCAGCTGGAGCTGCCACGCCCCCAAGAAATACAGTGCAGGAAAGTGTGCCCGAG 480
Db 677 CCCAGCTGGAGCTGCCACGCCCCCAAGAAATACAGTGCAGGAAAGTGTGCCCGAG 736
Qy 481 TGGGTATGTGACAGGGAGTGA---CACCGCGATCCAGCGCTCCACGGCGCAAGAC 537
Db 737 TGGGTATGTGACAGGGAGTGA---CACCGCGATCCAGCGCTCCACGGCGCAAGAC 796
Qy 538 CAATTTCTGCCCTTGTCTACTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 597
Db 797 CAATTTCTGCCCTTGTCTACTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 856
Qy 598 GCGTGGGCGCCCTGTCTCAACCACTGTGGCTGGGCAATGCCACCGAGTATCCAAACAG 657
Db 857 GCGTGGGCGCCCTGTCTCAACCACTGTGGCTGGGCAATGCCACCGAGTATCCAAACAG 916
Qy 658 AACGATTTCTGCACTGAGATGCAAGCGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 717
Db 917 AACGATTTCTGCACTGAGATGCAAGCGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 976
Qy 718 GCCAGAGCCACAGCTCATGTGAACAGTCTTTCTA 752
Db 977 TCCAGAGCCACAGGCTCATGTGAACAGTCTTTCTA 1001

```

## RESULT 2

US-09-182-145-18/c  
Sequence 18, Application US/09182145B  
Patent No. 6387657

## GENERAL INFORMATION:

APPLICANT: Botstein, David A.  
APPLICANT: Cohen, Robert  
APPLICANT: Goddard, Audrey  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: Lawrence, David A.  
APPLICANT: Levine, Arnold J.  
APPLICANT: Pennica, Diane  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME  
FILE REFERENCE: P1176R2  
CURRENT APPLICATION NUMBER: US/09/182,145B  
EARLIER FILING DATE: 1998-10-29  
EARLIER APPLICATION NUMBER: US 60/063,704  
EARLIER FILING DATE: 1997-10-29  
EARLIER APPLICATION NUMBER: US 60/073,612  
EARLIER FILING DATE: 1998-02-04  
EARLIER APPLICATION NUMBER: US 60/081,695  
EARLIER FILING DATE: 1998-04-14  
NUMBER OF SEQ ID NOS: 156  
SEQ ID NO 18  
LENGTH: 1734  
TYPE: DNA  
ORGANISM: Mus musculus  
US-09-182-145-18

Query Match 87.5%; Score 659; DB 4; Length 1734;  
Best Local Similarity 93.0%; Pred. No. 8.5e-164

```

Matches 702; Conservative 0; Mismatches 50; Indels 3; Gaps 1;
Qy 1 ATGAGGGGACCCAGCTGATCCATCTTGTGGCCACTTCTCTCTGCTTCTCTCATG 60
Db 1478 ATGAGGGGACCCAGCTGATCCATCTTGTGGCCACTTCTCTCTGCTTCTCTCATG 1419
Qy 61 GTGTGTGCCCACTGTGTGCGGACACCTGTACCTGTCTTGGACACACCCAGTGGCCA 120
Db 1418 GTGTATTCACAGCTGTGTGCGGACACCTGTCTCTCTTGTGGACACCCAGTGGCCA 1359
Qy 121 CAGGGGGTACCCCTGTGCTGTGATGGCTGTGGCTGTGCTGTGCTGTGCTGTGCTG 180
Db 1358 CCGGGGGTACCCCTGTGCTGTGATGGCTGTGGCTGTGCTGTGCTGTGCTGTGCTG 1299
Qy 181 GGGGAGTCTCTGACACACCTCATGTCTGACACCCACAGGCGCTGTGTGTACAGCT 240
Db 1298 GGGGAGTCTCTGACACACCTCATGTCTGACACCCACAGGCGCTGTGTGTACAGCT 1239
Qy 241 GGGGAGGCGCTGTGCGGCAATGGGGCTGTGTCTCTTGGATGAGATGAGGTACTGT 300
Db 1238 GGGGAGGCGCTGTGCGGCAATGGGGCTGTGTCTCTTGGATGAGATGAGGTACTGT 1179
Qy 301 GAGGTGATGGCCGAGGTACTGTGATGAGAGACCTTAAACCAATTGACAGGTCTG 360
Db 1178 GAGGTGATGGCCGAGGTACTGTGATGAGAGACCTTAAACCAATTGACAGGTCTG 1119
Qy 361 TGCCGCTGTGATGAGGTGTGCTTCACTGTGCTGCTGTGCTGCTGCTGCTGCTGCTG 420
Db 1118 TGCCGCTGTGATGAGGTGTGCTTCACTGTGCTGCTGTGCTGCTGCTGCTGCTGCTG 1059
Qy 421 CCCAGTGTGAGTGTGCGGCGCCCAAGAGATACAGGTGCGCAAGAAATGTGCGCGAG 480
Db 1058 CCCAGTGTGAGTGTGCGGCGCCCAAGAGATACAGGTGCGCAAGAAATGTGCGCGAG 999
Qy 481 TGGGTATGTGACAGGGAGTGA---CACCGCGATCCAGCGCTCCACGGCGCAAGAC 537
Db 998 TGGGTATGTGACAGGGAGTGA---CACCGCGATCCAGCGCTCCACGGCGCAAGAC 939
Qy 538 CAATTTCTGCCCTTGTCTACTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 597
Db 938 CAATTTCTGCCCTTGTCTACTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 879
Qy 598 GCGTGGGCGCCCTGTCTCAACCACTGTGGCTGGGCAATGCCACCGAGTATCCAAACAG 657
Db 878 GCGTGGGCGCCCTGTCTCAACCACTGTGGCTGGGCAATGCCACCGAGTATCCAAACAG 819
Qy 658 AACGATTTCTGCACTGAGATGCAAGCGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 717
Db 818 AACGATTTCTGCACTGAGATGCAAGCGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 759
Qy 718 GCCAGAGCCACAGCTCATGTGAACAGTCTTTCTA 752
Db 758 TCCAGAGCCACAGGCTCATGTGAACAGTCTTTCTA 724

```

## RESULT 3

US-09-182-145-13  
Sequence 13, Application US/09182145B  
Patent No. 6387657

## GENERAL INFORMATION:

APPLICANT: Botstein, David A.  
APPLICANT: Cohen, Robert  
APPLICANT: Goddard, Audrey  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: Lawrence, David A.  
APPLICANT: Levine, Arnold J.  
APPLICANT: Pennica, Diane  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME  
FILE REFERENCE: P1176R2  
CURRENT APPLICATION NUMBER: US/09/182,145B

```

: CURRENT FILING DATE: 1998-10-29
: EARLIER APPLICATION NUMBER: US 60/063,704
: EARLIER FILING DATE: 1997-10-29
: EARLIER APPLICATION NUMBER: US 60/073,612
: EARLIER FILING DATE: 1998-02-04
: EARLIER APPLICATION NUMBER: US 60/081,655
: EARLIER FILING DATE: 1998-04-14
: NUMBER OF SEQ ID NOS: 156
: SEQ ID NO 13
: LENGTH: 1293
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-09-182-145-13

```

|                           |       |                    |          |             |
|---------------------------|-------|--------------------|----------|-------------|
| Query Match               | 67.8% | Score 510.4        | DB 4     | Length 1293 |
| Best Local Similarity     | 79.9% | Pred. No. 1.4e-122 |          |             |
| Matches 601, Conservative | 0     | Mismatches 151     | Indels 0 | Gaps 0      |

|    |     |   |     |
|----|-----|---|-----|
| OY |     | ATGAGGGAGGCCACCTATTCATTCTTGTGGCACTCTTCCTTCCTGGCCTCTCTCAATG        | 60  |
| Dd | 22  | ATGAGAGGCACACCAGAAAGACCACCTCTGGCCCTTCTCCCTCTCTGTGCTCTTCAAAG       | 81  |
| OY | 61  | GTTGTGTCCCAAGCTGTGTCCGGACACCCTGTACTCTCTTGAGACACACCACAATGCCA       | 120 |
| Dd | 82  | GTCGTAACCAAGCTGTGTGCCGAACACATGTACTCTGCCCTGTGGCACCTCCCCGATGCCG     | 141 |
| OY | 121 | CAGGGGGTACCCCTGTGTCTGATGTAGTGTGTGCTCTCTTAAAGTGTGTGACGGAGCTG       | 180 |
| Dd | 142 | CTGGGGATACCCCTGTGTCTGTGATGTGTGTGTGTCTCTCGGGATGTATGTGCACGGCGCTG    | 201 |
| OY | 181 | GGGAGAGCCCGACACCACCTGCATGTCTGTCCAGACCCCAGACGGGCGTGTGTGCACCT       | 240 |
| Dd | 202 | GGGGAGCCCTGTGCACCAACTCCAAGCTGTGCAGCGCACGCCAGAGGCGCTGTGTGCACCCC    | 261 |
| OY | 241 | GGGGCAGGCCCTGTGGCGCCATGTGGGCTGTGTCTCTTGTGAATGAGAGTAGAGTAGCTGT     | 300 |
| Dd | 262 | GGGGCAGAACCCGGTGTGGCCGGGGGGGCCGTGTGCTCTTGGCAGAGAGACAGACAGCTGT     | 321 |
| OY | 301 | GAGGTGAATGTGCCGACAGTACTGTGATGTGAGAGACCTTTAAACCAATTCAGAGGCTCTG     | 360 |
| Dd | 322 | GAGGTGAACGGCCGCTGTATGTGGGAGAGGGGAGACCTTCCAGCCCCACTGCAGCATCTCGC    | 381 |
| OY | 361 | TGCGCGTGTGATAGCTGTGGCTTCAACCTGCGCGCGGTGTGCAGTGAAGAGATGTCCGCTG     | 420 |
| Dd | 382 | TGCGCGTGTGAGAGACGGCGGCTTCACTGTGCGCGCTGTGCAGAGAGAGATGTCCGCTG       | 441 |
| OY | 421 | CCCAGCTGGGACTGCCACGCGCCCAAGAAGATAACAGTGTGCAGAAAATGTCTGCCCGAG      | 480 |
| Dd | 442 | CCAGCGTGGGACTGTCCCCCACCAGAGAGGTGTGAGGTCTGTGGGCAMTGTCTGCCCTGAG     | 501 |
| OY | 481 | TGGGTATGTACCAAGGAGTGAACCCGGCGATCCAGCGCTCCAGCGCGCAGAGACACAA        | 540 |
| Dd | 502 | TGGGTGTGGCGCCCAAGAGGGGAGACTGTGGGAGACCCAGCCCTTCAGCCCAAGAGACCCAG    | 561 |
| OY | 541 | CTTTTCGCCCTTGTACTCTGTGCCTGTGTGTATGTCTCTTGTCCAAATTTGGAGACAGCC      | 600 |
| Dd | 562 | TTTTTCGTGGCTTGTCTTCTCCCTGTGCCCTCTGTGTGCCCTCTCCAGAAATGTGAGACGGCC   | 621 |
| OY | 601 | TGGGGGCCCTGTCAACACCACTGTGTGGGTGGGSCATAGGACCCGAGTGTCCAAACAGAAC     | 660 |
| Dd | 622 | TGGGAGACCTGTGTCAACCACTGTGTGGGTGTGGGSCATAGGACCCCGGGTGTCCAAACAGAAC  | 681 |
| OY | 661 | CGATTGTGCCAATGTGAGATCCAAAGCGCGCTGTGTGTGCCAGAACCCCTGTGCTGGACGC     | 720 |
| Dd | 682 | CGCTTGTGCCAATGTGAGAACCAAGCGCGCGCTGTGTGTGTCCAGGACCCCTGTGCCACCCCTCC | 741 |
| OY | 721 | AGGAGCACACAGCTCAATGAGAACAGTCTTTCTA                                | 752 |
| Dd | 742 | AGGGGTGTGCAATGTCAAAAAACAGTGTCTTTCA                                | 773 |

```

: Sequence 14, Application US/09182145B
: Patent No. 6387657
:
: GENERAL INFORMATION:
: APPLICANT: Botstein, David A.
: APPLICANT: Cohen, Robert
: APPLICANT: Goddard, Audrey
: APPLICANT: Gurney, Austin L.
: APPLICANT: Hillan, Kenneth J.
: APPLICANT: Lawrence, David A.
: APPLICANT: Levine, Arnold J.
: APPLICANT: Pennica, Diane
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Wood, William I.
: TITLE OF INVENTION: MISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
: FILE REFERENCE: P1176R2
: CURRENT APPLICATION NUMBER: US/09/182,145B
: CURRENT FILING DATE: 1998-10-29
: EARLIER APPLICATION NUMBER: US 60/063,704
: EARLIER FILING DATE: 1997-10-29
: EARLIER APPLICATION NUMBER: US 60/073,612
: EARLIER FILING DATE: 1998-02-04
: EARLIER APPLICATION NUMBER: US 60/081,695
: EARLIER FILING DATE: 1998-04-14
: NUMBER OF SEQ ID NOS: 156
:
: SEQ ID NO 14
: LENGTH: 1293
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-09-182-145-14

```

|                       |              |                   |                |             |
|-----------------------|--------------|-------------------|----------------|-------------|
| Query Match           | 67.8%        | Score 510.4       | DB 4           | Length 1293 |
| Best Local Similarity | 79.9%        | Pred No. 1.4e-122 |                |             |
| Matches 601           | Conservative | 0                 | Mismatches 151 | Indels 0    |
|                       |              |                   |                | Gaps 0      |

|    |      |   |       |
|----|------|---|-------|
| QY | 1    | TTGAGGGGACCCACCTCATTTCTTGGGACATTTCTTCCTTCCTGGCTTCTTCATG           | 60    |
| Db | 1272 | ATGAGAGGACACACGAGAACCCACCTCTCTGGCTTCTTCCTTCCTTCCTCTCAAG           | 12131 |
| QY | 61   | GTGTGTGCCAGCTGTGTGCCGAGACACCCCTGTACTGTCTTGGACACACCCGAGCCCA        | 120   |
| Db | 1212 | TGTGCTACCCAGCTGTGTGCCGACACCACTGTACTGTCTCTCCCTGGCCACTTCCCGATGGCCG  | 1153  |
| QY | 121  | CAGGGGGTACCCCTGTGTCTGTGATGTGGCTGTGGCTCTTAAAGTGTGTGCACGGAGCTG      | 180   |
| Db | 1152 | CTGGGAGTACCCCTGTGTCTGTGATGTGGCTGTGGCTCTCCGGGTATGTGCACGGGGCTG      | 1093  |
| QY | 181  | GGGGAGTCTGTGCAGACCTCGATGTGTGCAGCCCCAGCCACGAGGGCTGTGTTGTAGGCT      | 240   |
| Db | 1092 | GGGGAGCCCTGTGCACCACTCCACTCTGTGCAGCCGACGACGAGGGCTGTGTCCAGGCC       | 1033  |
| QY | 241  | GGGGGAGCCCTGTGGGCGCATGTGGGCTTGTGTCTTCTTGGATGTAGAGATGACGGTAGCTGT   | 300   |
| Db | 1032 | GGGGGAGGACCCGGGTGGGCGGGGGGCCCTGTGTGCTCTTGTGGAGAGGAGCAGCAGCGTGT    | 973   |
| QY | 301  | GAGTGAATGGCCGACGTTACCTGATGTGAGAGACCTTTAAACCAATTGCAGGGTCTCTG       | 360   |
| Db | 972  | GAGGTGAACGGCCGCTGTATCTGTGAGAAAGGGGAGACCTTCCAGCCCACTGTGACGCAATCCGC | 913   |
| QY | 361  | TGCCGCTGTATGATGAGTGTGGCTTACCGTCGCGCGCGGTGTGTGAGTGAAGATGTGGCGCTG   | 420   |
| Db | 912  | TGCCCTCTCGAGGAGGGCGGCTTACCTGTGCGCGCTGTGTGAGTGAAGATGTGGCGCTG       | 853   |
| QY | 421  | CCGAGCTGGAGCTGCCACGACCCCAAGAGAAATACAGTGTGCCAGAAATGTGTCGCCGAG      | 480   |
| Db | 852  | CCGAGCTGGAGCTGCCACCCCAAGAGAGGTCTGTGGCAAAATGTGTCGCCCTGAG           | 793   |
| QY | 481  | TGGGTATGTACCGAGGAGTGAACACCGGCGATCCAGCGCTCCACGGCGCAAGAGACCA        | 540   |
| Db | 792  | TGGGTGTGGGCGCAAGGAGGAGGAGTGTGGGAGCCGACCCCTTCCAGGCCCAAGAGACCCAG    | 733   |
| QY | 541  | CTTTCGTGCCCTTGTACCTGTGCTCTGTCTATGTCTCTTGTCCAAATTTGAGCAGACCC       | 600   |
| Db | 732  | TTTTCGTGCCCTTGTCTTCCCTGTGCCCTGTGTGTCCTCTGCCAGATGTAGACAGCGCC       | 673   |

```
QY 601 TGGGGCCCTGCTCAACCACTGTGGGCTGGGATAGCCACCCGAGTGTCCACAGAAC 660
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 672 TGGGGACCTCTTCACCACTGTGGGCTGGGATAGCCACCCGAGTGTCCACAGAAC 613
QY 661 CGATCTCCCACTGAGATCAACGCCGCTGTGTGGCCAGACGCTGGCTGGCAGCC 720
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 612 CGCTTCTCCGACGTGAACCCAGCGCCGCTGTGTGTCCAGGCGCTGCCACCTCC 553
QY 721 AGAGCCACAGCTCATGGAAGTGTCTTCTA 752
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 552 AGGGGTCCAGTCCACAAAGAGTGTCTTCTA 521

RESULT 5
US-09-182-145-38
; Sequence 38, Application US/09182145B
; Patent No. 6387657
; GENERAL INFORMATION:
; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/09/182,145B
; CURRENT FILING DATE: 1998-10-29
; EARLIER APPLICATION NUMBER: US 60/063,704
; EARLIER FILING DATE: 1997-10-29
; EARLIER APPLICATION NUMBER: US 60/073,612
; EARLIER FILING DATE: 1998-02-04
; EARLIER APPLICATION NUMBER: US 60/081,695
; EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 38
; LENGTH: 738
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-182-145-38

Query Match      66.6%; Score 501.2; DB 4; Length 738;
Best Local Similarity 79.9%; Pred. No. 2,9e-120;
Matches 590; Conservative 0; Mismatches 148; Indels 0; Gaps 0;

QY 13 CCACTGATCATCTTGTGGCACTTCTCTGCTGTCTCAATGATGTGCCAG 72
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 CCGAAGACCACTCTGCTCTCTCTCTCTCTCTCAAGTGTGCCAG 60
QY 73 CTGTGCGGACACCTGTACCTGTGTGGACACCCGAGTGTGCCACAGGGGTACC 132
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 CTGTGCGGACACCACTGTACCTGTGTGGACACCCGAGTGTGCCAGTGTGCC 120
QY 133 CTGTGCTGATGCTGTGGCTGTGTAAAGTGTGTGACAGGAGCTGGGAGTCTGC 192
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 CTGTGCTGATGCTGTGGCTGTGTAAAGTGTGTGACAGGAGCTGGGAGTCTGC 180
QY 193 GACCACTGATGTGTGGACCCAGCCAGGCGCTGTGTGTGTGTGTGTGTGTGT 252
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 GACCACTGATGTGTGGACCCAGCCAGGCGCTGTGTGTGTGTGTGTGTGTGT 240
QY 253 GCGGCGCATGGGCTGTGTCTTGTGATGAGATGACGAGTGTGAGTGTGATGAGC 312
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 GGTGGCGGGGGGCTGTGTCTTGTGAGAGGAGCAGACAGCTGTGAGTGTGAGCGC 300
QY 313 CGCAGTCTGATGTGAGAGACCTTTAAACCAATGTGAGGAGTCTGTGCGCTGTGAT 372
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 CGCTGTATCGGGAAGGAGACCTTCCAGGCCCACTGACAGATCGCTGCCGTGCGAG 360
```

```
QY 373 GACGCTGCTTACCTTGTGCTGCTGCTGTGCACTAGATGTGGGCTGCCAGCTGGAC 432
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 361 GACGCGGCTTACCTTGTGCTGCTGCTGTGCACTAGATGTGGGCTGCCAGCTGGAC 420
QY 433 TGCCGACGCCCAAGAGATGAGTGTGCAAGAAAGTCTGCCGAGTGGGTATGTGAC 492
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 421 TGCCGACGCCCAAGAGATGAGTGTGCAAGAAAGTCTGCCGAGTGGGTATGTGAC 480
QY 493 CAGGAGTGTGACCGGCGATCCAGCGCTTCACAGCGGCGCAAGACAACTTGTGCGCTT 552
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 481 CAGGAGGAGGAGCTGTGGACCAAGCCCTTCCAGGCGCAAGACCACTTGTGCGCTT 540
QY 553 GTCACTCTGCTTGTGATGCTCTTGTGCAATTTGAGACACAGCTGGGCGCTTGC 612
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 541 GTCTTCTCTGCGCCCTGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
QY 613 TCACACACCTGTGGGCTGGGATGACCAAGAGTGTCCACCAAGCAATTTGTGCCAA 672
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 601 TGACACCACTGTGGGCTGGGATGACCAAGAGTGTCCACCAAGCAATTTGTGCCAA 660
QY 673 CTGAGATTCACACCGCGCTGTGTGTGCTGCCAGACCTGCTGTGCCAGGAGGACACAGC 732
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 661 CTGAGACCAACAGCGCGCTGTGTGTGCTGCCAGACCTGCTGTGCCAGGAGGAGTGTG 720
QY 733 TCATGGAACAGTGTCTTC 750
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 721 CCACAAACAGTGTCTTC 738

RESULT 6
US-09-182-145-39
; Sequence 39, Application US/09182145B
; Patent No. 6387657
; GENERAL INFORMATION:
; APPLICANT: Botstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/09/182,145B
; CURRENT FILING DATE: 1998-10-29
; EARLIER APPLICATION NUMBER: US 60/063,704
; EARLIER FILING DATE: 1997-10-29
; EARLIER APPLICATION NUMBER: US 60/073,612
; EARLIER FILING DATE: 1998-02-04
; EARLIER APPLICATION NUMBER: US 60/081,695
; EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 39
; LENGTH: 841
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1-841
; OTHER INFORMATION: Sequence is synthesized.
; Patent No. 6387657
US-09-182-145-39

Query Match      65.9%; Score 496.4; DB 4; Length 841;
Best Local Similarity 79.7%; Pred. No. 5.2e-119;
Matches 599; Conservative 0; Mismatches 151; Indels 2; Gaps 1;

QY 1 ATGAGGGGACGCCCATCATCTTGTGCGCACTTCTCTCTCTCTCTCTCTCTCTCATG 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```



```

Db 12 ATGAGAGGACACCGAAGACCCACCTCTGCGCTTCTCCCTCTGCTCTCTCAAG 71
QY 61 GGTGTGCGCAGCTGTGCGGACACCTGTACTGTCTGTGAGACACACCCAGTCCCA 120
Db 72 GTGCTGACCGAGCTGTGCGGACACCTGTACTGTCTGCGGACCTCCCGATCCCG 131
QY 121 CAGGGGTACCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
Db 132 CTGGAGATACCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 191
QY 181 GGGGAGTCTGTGACACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 240
Db 192 GGGGAGTCTGTGACACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 251
QY 241 GGGGAGTCTGTGACACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 300
Db 252 GGGGAGTCTGTGACACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 311
QY 301 GAGGTGAATGCGCGACGAGTACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 360
Db 312 GAGGTGAATGCGCGACGAGTACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 371
QY 361 TGGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 420
Db 372 TGGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 431
QY 421 CCCAGCTGGGACGCGCCACGCGCCACGAGAAATACAGTGTGTGTGTGTGTGTGTGTGT 480
Db 432 CCCAGCTGGGACGCGCCACGCGCCACGAGAGGTGTGTGTGTGTGTGTGTGTGTGTGT 491
QY 481 TGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 540
Db 492 TGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 549
QY 541 CTCTTGTGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
Db 550 TTTTCTGTGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 609
QY 601 TGGGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660
Db 610 TGGGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 669
QY 661 CGATTCTGTGCAATGTGAGATGTCAACGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGT 720
Db 670 CGATTCTGTGCAATGTGAGATGTCAACGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGT 729
QY 721 AGGAGCCACAGCTCATGTGAGACAGTGTCTTCTA 752
Db 730 AGGAGCCACAGCTCATGTGAGACAGTGTCTTCTA 761

```

US-08-167-628-1  
Sequence 1, Application US/08167628  
Patent No. 5408040

## GENERAL INFORMATION:

APPLICANT: Grotenhorst, Gary R.  
APPLICANT: Bradham Jr., Douglas M.  
TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Spensley Horn Judas & Lubitz  
STREET: 4225 Executive Square, Suite 1400  
CITY: La Jolla  
STATE: CA  
COUNTRY: US  
ZIP: 92037

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:

```

APPLICATION NUMBER: US/08/167, 628
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/752, 427
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Wetherell, Jr., Ph.D., John W.
REGISTRATION NUMBER: 31,678
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-455-5100
TELEFAX: 619-455-5110
INFORMATION FOR SEQ. ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2075 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
IMMEDIATE SOURCE:
CLONE: DB60R32
FEATURE:
NAME/KEY: CDS
LOCATION: 130..1177
US-08-167-628-1

```

Query Match 21.6%; Score 163; DB 1; Length 2075;  
Best Local Similarity 56.2%; Pred. No. 4.5e-33;  
Matches 351; Conservative 0; Mismatches 265; Indels 9; Gaps 2;

```

QY 102 GACACACCCAGTGGCCACAGAGGGGTACCCTGTGTGTGTGTGTGTGTGTGTGTGTGT 161
Db 243 GCGGCGCGCGCGCTGCGCGCGCGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 302
QY 162 AGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 221
Db 303 CGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 362
QY 222 GGGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 281
Db 363 GGGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 419
QY 282 TGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 341
Db 420 CAAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 479
QY 342 ACCCAATTGCAGAGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 401
Db 480 GAGCAGCTGTCAAGTACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 539
QY 402 CAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 461
Db 540 CAGCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 599
QY 462 AGGAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 521
Db 600 CCGGAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 659
QY 522 CACGGCGAAGGACACCACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 575
Db 660 TGGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 719
QY 576 TCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 635
Db 720 CAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 779
QY 636 AGCAGCCGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 695
Db 780 CTCACCCGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 839
QY 696 TCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 720

```





Db 720 CAATGCTGCTCCAGACAGAGTGAGCGCCCTGTTCCAGACCTGTGGGATGGCAT 779  
QY 636 AGCCACCCGAGTGTCCACAGAACCGATTCTGCACTGAGATCCAAACCGCCCTGTG 695  
Db 780 CTCACCGCGGTTACATGACAAAGCGCTCTGAGAGGTAGAGAAGCAGACCGCCCTGTG 839  
QY 696 TCTGCCAGACCTTGCCTGGCAGCC 720  
Db 840 CATGTCAGCGCTTGCAGAGCTGAC 864

## RESULT 11

US-08-880-031-1  
; Sequence 1, Application US/08880031  
; Patent No. 5916756  
; GENERAL INFORMATION:  
; APPLICANT: Grotenordt, Gary R.  
; APPLICANT: Bradham Jr., Douglas M.  
; TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Spensley Horn Judas & Lubitz  
; STREET: 4225 Executive Square, Suite 1400  
; CITY: La Jolla  
; STATE: CA  
; COUNTRY: US  
; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/880.031  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/167.628  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Wetherell, Jr. Ph.D., John W.  
; REGISTRATION NUMBER: 31.678  
; REFERENCE/DOCKET NUMBER: PD-1294  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619-455-5100  
; TELEFAX: 619-455-5110  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2075 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; IMMEDIATE SOURCE:  
; CLONE: DB60R32  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 130..1177  
; US-08-880-031-1

Query Match 21.6%; Score 163; DB 2; Length 2075;  
Best Local Similarity 36.2%; Pred. No. 4.5e-33;  
Matches 351; Conservative 0; Mismatches 265; Indels 9; Gaps 2;

QY 102 GACACACCCAGCCAGCCAGAGGGGTACCCCTGCTGCTGATGGCTGGCTGCTGTA 161  
Db 243 GCGGCGCGCGCTGCGCGCGCGCGCGCTGAGCTTCTGCTGACGCGCTGCGCTGCGG 302  
QY 162 AGTGTGTGACGAGAGCTGGGGAGTCTTGGCAGCCAGCTGATGTCTGAGACCCAGCA 221  
Db 303 CGTGTGGCGCAAGAGCTGGGGAGCTGTGACAGAGGGAGCCCTGAGACCCAGCA 362  
QY 222 GGGCGCTGTTGTCAAGCTGGGGAGGCGCCCTGGCGCCAGTGGGGCTGTGTCTTGA 281

Db 363 GGGCTCTTCTGTAGCTTTCGCTCCCGGCGCAACGCAAGATCGCGTGTG---CACCC 419  
QY 282 TGAGGATGACGGTATGCTGTGAGTGAATGCGCGAGATACCTGGATGAGAGACTTAA 341  
Db 420 CAAGATGTGCTCTCCCTGACATCTTCCGTGTGAGGTGTACCGCGAGAGAGTCTTCCA 479  
QY 342 ACCCAATTGACAGGCTCTGTCGCCCTGTGATGACGCTTGCTTACCTGCTCCGCTGTG 401  
Db 480 GAGCAGCTGCAAGTACAGTGTACAGTGTGAGCGGGCGGTGGCTGATGATGATGATG 539  
QY 402 CAGTGGATGTGGGGCTGCGCCAGCTGGAGTGGCCAGCCCAAGAAATATACAGTACC 461  
Db 540 CAGCATGAGCTTGTGTCTGCTGCCAGCTGATGCTGCTTCCGAGAGAGGTCAAGCTGCC 599  
QY 462 AGGAAGTGTGCCCCGAGGGGTATGTGACAGAGGAGTGAACACCGGATTCAGCGCTC 521  
Db 600 CGGAAATGTCTGCGAGAGGTGTGTGTGAGAGAGCCCAAGAACCCGTGTTGGGCC 659  
QY 522 CACGCGCGAAGACACCACTTTTGCCTTGTCACTCTCTCTCTGTGAT-----GC 575  
Db 660 TGCCCTCGCGCTTACCGACTGGAAGACAGCTTTGGCCCAAGACCACTATGATTAAGC 719  
QY 576 TCTTGTCCAAATTGAGACAGAGCTGGGGCCCTGCTCAACACCACTGTGGCTGGCAT 635  
Db 720 CAATGCTGCTCCAGACACAGAGTGGAGCGCTGTTCAGAGACTGTGGATGGCAT 779  
QY 636 AGCCACCCGAGTGTCCACAGAACCGATTCTGCCAAGTGAATCCAAACCGCCCTGTG 695  
Db 780 CTCACCGCGGTTACATGACAAAGCGCTCTGAGGTTAGAAAGCAGACCGCCCTGTG 839  
QY 696 TCTGCCAGACCTTGCCTGGCAGCC 720  
Db 840 CATGTCAGCGCTTGCAGAGCTGAC 864

## RESULT 12

US-09-097-179-1  
; Sequence 1, Application US/09097179  
; Patent No. 6149916  
; GENERAL INFORMATION:  
; APPLICANT: Grotenordt, Gary R.  
; APPLICANT: Bradham Jr., Douglas M.  
; TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Spensley Horn Judas & Lubitz  
; STREET: 4225 Executive Square, Suite 1400  
; CITY: La Jolla  
; STATE: CA  
; COUNTRY: US  
; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/097.179  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/386.680  
; FILING DATE: 10-FEB-1995  
; APPLICATION NUMBER: US/08/167.628  
; FILING DATE:  
; APPLICATION NUMBER: US/07/752.427  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Wetherell, Jr. Ph.D., John W.  
; REGISTRATION NUMBER: 31.678  
; REFERENCE/DOCKET NUMBER: PD-1294  
; TELECOMMUNICATION INFORMATION:

TELEPHONE: 619-455-5100  
TELEFAX: 619-455-5110  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2075 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cdna  
IMMEDIATE SOURCE:  
CLONE: DB60R32  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 130..1177  
US-09-097-179-1

Query Match 21.6%; Score 163; DB 3; Length 2075;  
Best Local Similarity 56.2%; Pred. No. 4.5e-33;  
Matches 351; Conservative 0; Mismatches 265; Indels 9; Gaps 2;

102 GACACACACCCAGTCCACAGAGGGGTACCCCTGCTGGATGCTGTGCTCTGTAA 161  
243 GCCGGCCCGCCGCTGCGCCGCGGCGGTGAGCTCTGTGACGCGCTGCGCTGCG 302  
162 AGTGTGTGACAGGAGCTGGGGAGTCTGTGACACCTGCTGTGACACCCAGCA 221  
303 CGTGTGGCCAGAGCAGTGGGGAGTGTGACACCGAGCCCTGTGACCCGACAA 362  
222 GGGCTGTGTGTGACCTGTGGGAGGCGCTGGCGGCGCATGGGCTGTGTCTTGA 281  
363 GGGCTGTGTGTGACCTGTGGGAGGCGCTGGCGGCGCATGGGCTGTGTCTTGA 419  
282 TGAGATGAGCGTGTGTGAGTGAATGGCGGCGGAGTGTGAGTGAATGAGAGAT 341  
420 CAAGATGAGTGTGTGTGATCTTCTGCTGTGATGAGTGTGATGAGTGTGATG 479  
342 ACCCAATTGAGGCTGTGTGCTGTGATGAGTGTGATGAGTGTGATGAGTGTG 401  
480 GAGCAGCTGTGATGAGTGTGATGAGTGTGATGAGTGTGATGAGTGTGATG 539  
402 CAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 461  
540 CAGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 599  
462 AGGAAATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 521  
600 CGGAAATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 659  
522 CACGGCGCAAGAGACACCACTTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 575  
660 TGCCCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 719  
576 TCCCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 635  
720 CAATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 779  
636 AGCAGCGAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 695  
780 CTCACCGCGGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 839  
696 TCTGCGCAGACCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 720  
840 CATGCTGAGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 864

RESULT 13  
US-09-080-715-1  
Sequence 1, Application us/09080715  
Patent No. 6190884  
GENERAL INFORMATION:  
APPLICANT: Grotendorf, Gary R.  
APPLICANT: Bradham Jr., Douglas M.  
TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR

NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Spensley Horn Jubas & Lubitz  
STREET: 4225 Executive Square, Suite 1400  
CITY: La Jolla  
STATE: CA  
COUNTRY: US  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/080,715  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/167,628  
FILING DATE:  
APPLICATION NUMBER: US/07/752,427  
ATTORNEY/AGENT INFORMATION:  
NAME: Wetherell, Jr., Ph.D., John W.  
REGISTRATION NUMBER: 31,678  
REFERENCE/DOCKET NUMBER: PD-1294  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-455-5100  
TELEFAX: 619-455-5110

INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2075 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cdna  
IMMEDIATE SOURCE:  
CLONE: DB60R32  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 130..1177  
US-09-080-715-1

Query Match 21.6%; Score 163; DB 4; Length 2075;  
Best Local Similarity 56.2%; Pred. No. 4.5e-33;  
Matches 351; Conservative 0; Mismatches 265; Indels 9; Gaps 2;

102 GACACACACCCAGTCCACAGAGGGGTACCCCTGCTGGATGCTGTGCTCTGTAA 161  
243 GCCGGCCCGCCGCTGCGCCGCGGCGGTGAGCTCTGTGACGCGCTGCGCTGCG 302  
162 AGTGTGTGACAGGAGCTGGGGAGTCTGTGACACCTGCTGTGACACCCAGCA 221  
303 CGTGTGGCCAGAGCAGTGGGGAGTGTGACACCGAGCCCTGTGCGACCA 362  
222 GGGCTGTGTGTGACCTGTGGGAGGCGCTGGCGGCGCATGGGCTGTGTCTTGA 281  
363 GGGCTGTGTGTGACCTGTGGGAGGCGCTGGCGGCGCATGGGCTGTGTCTTGA 419  
282 TGAGATGAGCGTGTGTGAGTGAATGGCGGCGAGTGTGATGAGTGAATGAGAG 341  
420 CAAGATGAGTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 479  
342 ACCCAATTGAGGCTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 401  
480 GAGCAGCTGTGATGAGTGTGATGAGTGTGATGAGTGTGATGAGTGTGATGAG 539  
402 CAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 461  
540 CAGATGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 599  
462 AGGAAATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 521

US-09-080-715-1  
Sequence 1, Application us/09080715  
Patent No. 6190884  
GENERAL INFORMATION:  
APPLICANT: Grotendorf, Gary R.  
APPLICANT: Bradham Jr., Douglas M.  
TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR

Db 600 CGGGAATGCTGCGAGAGTGGTGTGACGAGCCCAAGACCAACCGTGGTGGCC 659  
QY 522 CACGGCCAGAGACCACTTTCTGCCCTTGTCACTCCCTCTCTGCTGAT-----GC 575  
Db 660 TGCCCTGGCGGCTTACCGACTGAAGACGTTTGGCCAGACCCCACTATGATTAGAGC 719  
QY 576 TCTTGTCCAAATTTGAGACACAGCCTGGGGCCCTGCTCAACCACTTGGGCTGGGAT 635  
Db 720 CAACCTGCTGTGTCCAGACACAGAGTGGAGCGCCTGTGGAAGACCTGGGATGGCAT 779  
QY 636 AGCCACCGAGTGTCCACAGACCGATCTGCCACTGGAGATCCAAACCGCGCTGTG 695  
Db 780 CTCACCCGGGTTACCAATGACAAAGCCTCTGCAAGGTTAGAGAGCAGACCGCTGTG 839  
QY 696 TCTGCCAGACCGCTGCTGGCAGCC 720  
Db 840 CATGTCAGGCTTGGCAAGCTGAC 864

RESULT 14  
US-09-142-569-7  
; Sequence 7, Application US/09142569  
; Patent No. 6413735  
; GENERAL INFORMATION:  
; APPLICANT: Lau, Lester F.  
; TITLE OF INVENTION: Extracellular Matrix Signalling Molecules  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States of America  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/142,569  
; FILING DATE: 02-Apr-1999  
; CLASSIFICATION: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Clough, David W.  
; REGISTRATION NUMBER: 36,107  
; REFERENCE/DOCKET NUMBER: 28758/33766  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/474-6300  
; TELEFAX: 312/474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2075 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: "CTGF cDNA coding sequence"  
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:  
US-09-142-569-7

Query Match 21.6%; Score 163; DB 4; Length 2075;  
Best Local Similarity 56.2%; Pred. No. 4.5e-33;  
Matches 351; Conservative 0; Mismatches 265; Indels 9; Gaps 2;

QY 102 GACACACACCCAGTGTCCACAGAGGGGTACCCCTGTGTGATGGCTGTGCTGTAA 161  
Db 243 GCCGGCGCGGCTGCGCGGGGGGTGAGCCTGTGCTGAGCGGCTGCGGCTGTGCGG 302  
QY 162 AGTGTGTACAGGAGGTGGGGAGTCTCTGGAGCACTGATGTCTGCGACCCAGCA 221

Db 303 CGTGTGCCCAAGAGAGCTGGGCGAGCTGTGACGAGGCGCAACCCCTGCGACCCGACAA 362  
QY 222 GGGCCTGTTGATGAGCCTTGGGGCAGAGCCCTGGGGCCATGGGGCTGTGCTCTTTGA 281  
Db 363 GGGCCTCTTGTGACTTGGGCTCCCGGCCAACCGCAAGATGCGCGTGTG---CAAGC 419  
QY 282 TGAGATGACGATGACTGTGATGATGATGATGATGATGATGATGATGATGATGAT 341  
Db 420 CAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 479  
QY 342 ACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 401  
Db 480 GAGCAGCTGCAAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 539  
QY 402 CAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 461  
Db 540 CAGCATGAGAGCTGTGCTGTGCGCCAGCGCTGACCTGCGCCCTTCCGAGAGAGGTCAAGCTGC 599  
QY 462 AGGAATGCTGTGCGCCAGTGGGTATGTGACCAAGGAGTACACCGCGGATCCAGCGTC 521  
Db 600 CGGGAATGCTGCGAGAGTGGGTGTGTGACGAGGCGCAAGGACCAACCGTGTGGCC 659  
QY 522 CACGGCCAGAGACCAACTTTCTGCCCTTGTCACTCCCTCTCTGCTGAT-----GC 575  
Db 660 TGCCCTGGCGGCTTACCGACTGAAGACAGCTTTGGCCAGACCCCACTATGATTAGAGC 719  
QY 576 TCTTGTCCAAATTTGAGACACAGCCTGGGGCCCTGCTCAACCACTTGGGCTGGGAT 635  
Db 720 CAACCTGCTGTGTCCAGACACAGAGTGGAGCGCCTGTGGAAGACCTGGGATGGCAT 779  
QY 636 AGCCACCGAGTGTCCACAGACCGATCTGCCACTGGAGATCCAAACCGCGCTGTG 695  
Db 780 CTCACCCGGGTTACCAATGACAAAGCCTCTGCAAGGTTAGAGAGCAGACCGCGCTGTG 839  
QY 696 TCTGCCAGACCGCTGCTGGCAGCC 720  
Db 840 CATGTCAGGCTTGGCAAGCTGAC 864

RESULT 15  
PCT-US96-08140-1  
; Sequence 1, Application PC/TUS9608140  
; GENERAL INFORMATION:  
; APPLICANT: University of South Florida  
; TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 4225 Executive Square, Suite 1400  
; CITY: La Jolla  
; STATE: CA  
; COUNTRY: US  
; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US96/08140  
; FILING DATE: 30-MAY-1996  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Haile, Ph.D., Lisa A.  
; REGISTRATION NUMBER: 38,347  
; REFERENCE/DOCKET NUMBER: 07414/003W01  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619-678-5070  
; TELEFAX: 619-678-5099  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2075 base pairs

TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
IMMEDIATE SOURCE:  
CLONE: CTGF  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 130..1177  
PCT-US96-08140-1

Query Match 21.6%; Score 163; DB 5; Length 2075;  
Best Local Similarity 56.2%; Pred. No. 4.5e-33;

Matches 351; Conservative 0; Mismatches 265; Indels 9; Gaps 2;

```
QY 102 GACACACCCCACTGCCCCACAGGGGTTACCCCTGTGTGATGCTGTGCTCTGTAA 161
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 243 GCCGGCCCGCGCTGCCCGCGCGCTGACCTCTGTGTGACGGCTGCGCTGCTCCG 302
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 162 AGTGTGTACACGAGGAGCTGGGAGTCTGCGACACCTGCATGTGCGACCCACGCA 221
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 303 CGTCTGGCCCAACAGCAGTGGGCGAGCTGTGACCGACGCGACCCCTGCGACCCGACA 362
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 222 GGGCCTGTGTTGTACGCTGGGCGAGGCGCTGCGGCGCATGGGCTGTGTCTTGA 281
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 363 GGGCCTGTGTTGTACGCTGGGCGAGGCGCTGCGGCGCATGGGCTGTGTCTTGA 419
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 282 TGAGGATGACGTTAGCTGTAGTGAATGGCCCGAGTACTGTGATGAGAGACTTAA 341
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 420 CAAGATGTGTCTCTGCACTTCTGCTGATGCTGATGCTGATGCTGATGCTGATGCT 479
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 342 ACCCAATTGACAGGTTCTGTGCGCTGTGATGACGTTGCTGCTGCTGCTGCTG 401
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 480 GACGAGTGAATGACAGTGTGACGCTGTGACGAGGCGGCTGCTGCTGCTGCTG 539
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 402 CAGTGAGATGTGGGCTGCCACGCTGGGACTGGCCACGCCCCAAGAGAAATACAGTGC 461
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 540 CAGCATGTGACGTTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 599
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 462 AGGAAGTGTGCTGCCCGAGGAGTGTGACGAGGATGACACCGGCGATCCAGCGCTC 521
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 600 CGGGAATGTGTGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 659
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 522 CACGGCGCAGACACCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 575
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 660 TGCCCTTGGCGCTTACCGACTGGAAGACAGTTTGGCCAGACCACTATGATTAGAG 719
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 576 TCCTTGTCAAATTTGAGACAGCTGGGCGCCCTGCTCAACCACTGTGGCTGGGAT 635
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 720 CAACGTGCTGTGCTGACAGCAGAGTGTGAGCGGCTGTTCCAGACCTGTGGATGGGAT 779
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 636 AGCCACCCGAGTGTGCCACACAGAACGATCTGCAACTGTGATCCACGCGGCTGTG 695
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 780 CTCACACCCGGGTTACCAATGATCAACAGGCTGCTGAGGCTGAGAAAGCAGAGCGGCTGTG 839
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 696 TCTGCCAGACCTGCTGCGACGC 720
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 840 CATGTCTAGGCTTTCGAAGCTGAC 864
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Search completed: July 28, 2003, 21:36:01  
Job time : 43.1673 secs





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 15:36:41 ; Search time 185.622 Seconds  
(without alignments)  
8368.853 Million cell updates/sec

Title: US-10-010-408-3

Perfect score: 753

Sequence: 1 ATGAGGGGCGCCACTGAT.....CATGAACAGTCTTCTAA 753

Scoring table: IDENTITY\_NMC

Searched: 1439767 seqs, 1031500376 residues

Total number of hits satisfying chosen parameters: 2879534

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: Published\_Applications\_NA:\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
10: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq1.\*  
11: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq2.\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq3.\*  
13: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
15: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*  
16: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*  
17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | ID                | Description        |
|------------|-------|-------------|--------|-------------------|--------------------|
| 1          | 753   | 100.0       | 753    | US-10-010-408-3   | Sequence 3, Appl1  |
| 2          | 753   | 100.0       | 1708   | US-10-010-408-1   | Sequence 1, Appl1  |
| 3          | 681   | 90.4        | 681    | US-10-010-408-12  | Sequence 12, Appl1 |
| 4          | 659   | 87.5        | 1734   | US-10-112-267-17  | Sequence 17, Appl1 |
| 5          | 659   | 87.5        | 1734   | US-10-112-267-18  | Sequence 18, Appl1 |
| 6          | 510.4 | 67.8        | 1266   | US-10-137-866-319 | Sequence 319, App  |
| 7          | 510.4 | 67.8        | 1266   | US-10-146-726-319 | Sequence 319, App  |
| 8          | 510.4 | 67.8        | 1266   | US-10-146-727-319 | Sequence 319, App  |
| 9          | 510.4 | 67.8        | 1266   | US-10-146-788-319 | Sequence 319, App  |
| 10         | 510.4 | 67.8        | 1266   | US-10-153-380-319 | Sequence 319, App  |
| 11         | 510.4 | 67.8        | 1266   | US-10-153-934-319 | Sequence 319, App  |
| 12         | 510.4 | 67.8        | 1266   | US-10-028-072-319 | Sequence 319, App  |
| 13         | 510.4 | 67.8        | 1266   | US-10-121-049-319 | Sequence 319, App  |
| 14         | 510.4 | 67.8        | 1266   | US-10-123-904-319 | Sequence 319, App  |
| 15         | 510.4 | 67.8        | 1266   | US-10-140-470-319 | Sequence 319, App  |
| 16         | 510.4 | 67.8        | 1266   | US-10-175-746-319 | Sequence 319, App  |

|    |       |      |      |                    |                   |
|----|-------|------|------|--------------------|-------------------|
| 17 | 510.4 | 67.8 | 1266 | US-10-176-918-319  | Sequence 319, App |
| 18 | 510.4 | 67.8 | 1266 | US-10-176-921-319  | Sequence 319, App |
| 19 | 510.4 | 67.8 | 1266 | US-10-137-865-319  | Sequence 319, App |
| 20 | 510.4 | 67.8 | 1266 | US-10-140-474-319  | Sequence 319, App |
| 21 | 510.4 | 67.8 | 1266 | US-10-142-431-319  | Sequence 319, App |
| 22 | 510.4 | 67.8 | 1266 | US-10-143-114-319  | Sequence 319, App |
| 23 | 510.4 | 67.8 | 1266 | US-10-140-002-319  | Sequence 319, App |
| 24 | 510.4 | 67.8 | 1266 | US-10-142-419-319  | Sequence 319, App |
| 25 | 510.4 | 67.8 | 1266 | US-10-123-262-319  | Sequence 319, App |
| 26 | 510.4 | 67.8 | 1266 | US-10-142-423-319  | Sequence 319, App |
| 27 | 510.4 | 67.8 | 1266 | US-10-121-050-319  | Sequence 319, App |
| 28 | 510.4 | 67.8 | 1266 | US-10-141-755-319  | Sequence 319, App |
| 29 | 510.4 | 67.8 | 1266 | US-10-143-032-319  | Sequence 319, App |
| 30 | 510.4 | 67.8 | 1266 | US-10-123-108-319  | Sequence 319, App |
| 31 | 510.4 | 67.8 | 1266 | US-10-123-236-319  | Sequence 319, App |
| 32 | 510.4 | 67.8 | 1266 | US-10-123-261-319  | Sequence 319, App |
| 33 | 510.4 | 67.8 | 1266 | US-10-140-921-319  | Sequence 319, App |
| 34 | 510.4 | 67.8 | 1266 | US-10-140-928-319  | Sequence 319, App |
| 35 | 510.4 | 67.8 | 1266 | US-10-121-045-319  | Sequence 319, App |
| 36 | 510.4 | 67.8 | 1266 | US-10-123-292-319  | Sequence 319, App |
| 37 | 510.4 | 67.8 | 1266 | US-10-123-903-319  | Sequence 319, App |
| 38 | 510.4 | 67.8 | 1266 | US-10-124-819-319  | Sequence 319, App |
| 39 | 510.4 | 67.8 | 1266 | US-10-124-822-319  | Sequence 319, App |
| 40 | 510.4 | 67.8 | 1266 | US-10-140-925-319  | Sequence 319, App |
| 41 | 510.4 | 67.8 | 1266 | US-10-160-498-319  | Sequence 319, App |
| 42 | 510.4 | 67.8 | 1266 | US-10-124-824-319  | Sequence 319, App |
| 43 | 510.4 | 67.8 | 1266 | US-10-127-825A-319 | Sequence 319, App |
| 44 | 510.4 | 67.8 | 1266 | US-10-127-829A-319 | Sequence 319, App |
| 45 | 510.4 | 67.8 | 1266 | US-10-127-835A-319 | Sequence 319, App |

## ALIGNMENTS

RESULT 1  
US-10-010-408-3  
Sequence 3, Application US/10010408  
Publication No. US20020165185A1  
GENERAL INFORMATION:  
APPLICANT: John J. Castelli, Jr.  
TITLE OF INVENTION: No. US20020165185A1 Heparin-Induced CCN-Like Molecules and Uses Therefor  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD, LLP  
STREET: 28 State Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/010,408  
FILING DATE: 07-Dec-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/044,273  
FILING DATE: March 19, 1998  
APPLICATION NUMBER: <Unknown>  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Amy E. Mandragouras  
REGISTRATION NUMBER: 36,207  
REFERENCE/DOCKET NUMBER: MBI-004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)742-4214  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:

LENGTH: 753 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cdna  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 1..750  
 SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
 US-10-010-408-3

Query Match 100.0%; Score 753; DB 15; Length 753;  
 Best Local Similarity 100.0%; Pred. No. 7,4e-206;  
 Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 ATGAGGGGACCCACCTGATCATCTTCTGCGCACTTCTCTCTGCTCTCTCATG 60
DB 1 ATGAGGGGACCCACCTGATCATCTTCTGCGCACTTCTCTCTGCTCTCTCATG 60
OY 61 GTGTGTGCCAGCTGTGTGCGGACACCCCTGATCTCTCTGAGACACCCAGTCCCA 120
DB 61 GTGTGTGCCAGCTGTGTGCGGACACCCCTGATCTCTCTGAGACACCCAGTCCCA 120
OY 121 CAGGGGGTACCCCTGTGTGCTGTGATGCTGTGTGTGTGTGTGTGTGTGTGTGT 180
DB 121 CAGGGGGTACCCCTGTGTGCTGTGATGCTGTGTGTGTGTGTGTGTGTGTGTGT 180
OY 181 GGGGAGTCTCTGACACCACTGATGTGTGCGACCCACAGGGGCTGTGTGTGTGTGT 240
DB 181 GGGGAGTCTCTGACACCACTGATGTGTGCGACCCACAGGGGCTGTGTGTGTGTGT 240
OY 241 GGGGAGGCTCTGACACCACTGATGTGTGCGACCCACAGGGGCTGTGTGTGTGTGT 300
DB 241 GGGGAGGCTCTGACACCACTGATGTGTGCGACCCACAGGGGCTGTGTGTGTGTGT 300
OY 301 GAGGTGAATGGCCCGACAGTACCTGATGATGAGAGACCTTAAACCAATTGAGGCTCTG 360
DB 301 GAGGTGAATGGCCCGACAGTACCTGATGATGAGAGACCTTAAACCAATTGAGGCTCTG 360
OY 361 TGCCGCTGTGATGACAGTGTGCTTACCTGCTGCTGTGTGTGTGTGTGTGTGTGTGT 420
DB 361 TGCCGCTGTGATGACAGTGTGCTTACCTGCTGCTGTGTGTGTGTGTGTGTGTGTGT 420
OY 421 CCCAGCTGTGACAGTGTGCTTACCTGCTGCTGTGTGTGTGTGTGTGTGTGTGTGT 480
DB 421 CCCAGCTGTGACAGTGTGCTTACCTGCTGCTGTGTGTGTGTGTGTGTGTGTGTGT 480
OY 481 TGGGTATGTGACAGGAGTGTGACACCGGCGATCCAGCGCTCCAGCGGCCAAGGACACCA 540
DB 481 TGGGTATGTGACAGGAGTGTGACACCGGCGATCCAGCGCTCCAGCGGCCAAGGACACCA 540
OY 541 CTTTCTGCGCTTGTACACTGTGCTTGTGATGCTCTTGTCCAAATTGGAGCACACC 600
DB 541 CTTTCTGCGCTTGTACACTGTGCTTGTGATGCTCTTGTCCAAATTGGAGCACACC 600
OY 601 TGGGGCCCTGCTCAACCACTGTGTGGCTGGGACATAGCCACCCAGTGTCCAAACAGAAC 660
DB 601 TGGGGCCCTGCTCAACCACTGTGTGGCTGGGACATAGCCACCCAGTGTCCAAACAGAAC 660
OY 661 CGATTTCGCCAAGTGAATGCAACGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 720
DB 661 CGATTTCGCCAAGTGAATGCAACGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 720
OY 721 AGGAGCCACAGCTCATGAGACAGTCTTTCTAA 753
DB 721 AGGAGCCACAGCTCATGAGACAGTCTTTCTAA 753

```

RESULT 2  
 US-10-010-408-1  
 : Sequence 1, Application US/10010408  
 : Publication No. US20020165185A1  
 : GENERAL INFORMATION:

APPLICANT: John J. Castellot, Jr.  
 TITLE OF INVENTION: No. US20020165185A1e1 Heparin-Induced CCN-Like Molecules  
 and Uses Therefor  
 NUMBER OF SEQUENCES: 13  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: LAHIVE & COCKFIELD, LLP  
 STREET: 28 State Street  
 CITY: Boston  
 STATE: Massachusetts  
 COUNTRY: USA  
 ZIP: 02109  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/10/010,408  
 FILING DATE: 07-Dec-2001  
 CLASSIFICATION: <unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 09/044,273  
 FILING DATE: March 19, 1998  
 APPLICATION NUMBER: <unknown>  
 FILING DATE: <unknown>  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Amy E. Mandragouras  
 REGISTRATION NUMBER: 36,207  
 REFERENCE/DOCKET NUMBER: MBI-004  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (617)227-7400  
 TELEFAX: (617)742-4214  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1708 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cdna  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 249..1001  
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
 US-10-010-408-1

Query Match 100.0%; Score 753; DB 15; Length 1708;  
 Best Local Similarity 100.0%; Pred. No. 8.2e-206;  
 Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 ATGAGGGGACCCACCTGATCATCTTCTGCGCACTTCTCTCTGCTCTCTCATG 60
DB 249 ATGAGGGGACCCACCTGATCATCTTCTGCGCACTTCTCTCTGCTCTCTCATG 308
OY 61 GTGTGTGCCAGCTGTGTGCGGACACCCCTGATCTCTCTGAGACACCCAGTCCCA 120
DB 309 GTGTGTGCCAGCTGTGTGCGGACACCCCTGATCTCTCTGAGACACCCAGTCCCA 368
OY 121 CAGGGGGTACCCCTGTGTGCTGTGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
DB 369 CAGGGGGTACCCCTGTGTGCTGTGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 428
OY 181 GGGGAGTCTCTGACACCACTGATGTGTGCGACCCACAGGGGCTGTGTGTGTGTGTGT 240
DB 429 GGGGAGTCTCTGACACCACTGATGTGTGCGACCCACAGGGGCTGTGTGTGTGTGTGT 488
OY 241 GGGGAGGCTCTGAGGCGCATGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 300
DB 489 GGGGAGGCTCTGAGGCGCATGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 548
OY 301 GAGGTGAATGGCCCGACAGTACCTGATGATGAGAGACCTTAAACCAATTGAGGCTCTG 360
DB 549 GAGGTGAATGGCCCGACAGTACCTGATGATGAGAGACCTTAAACCAATTGAGGCTCTG 608

```



```

1  APPLICANT: WOOD, WILLIAM I.
2  TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
3  FILE REFERENCE: P1176R2
4  CURRENT APPLICATION NUMBER: US/10/112,267
5  CURRENT FILING DATE: 2002-03-27
6  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/182,145B
7  PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-29
8  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/063,704
9  PRIOR FILING DATE: EARLIER FILING DATE: 1997-10-29
10 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/073,612
11 PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-04
12 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/081,695
13 NUMBER OF SEQ ID NOS: 156
14 SEQ ID NO 17
15 LENGTH: 1734
16 TYPE: DNA
17 ORGANISM: Mus musculus
18 US-10-112-267-17

```

|                       |              |           |               |             |
|-----------------------|--------------|-----------|---------------|-------------|
| Query Match           | 87.5%        | Score 659 | DB 15         | Length 1734 |
| Best Local Similarity | 93.0%        | Pred. NC  | 6.6e-179      |             |
| Matches 702           | Conservative | 0         | Mismatches 50 | Indels 3    |
|                       |              |           |               | Gaps 1      |

|    |     |  |     |
|----|-----|--|-----|
| QY | 1   | ATGAGGGGCAACCCACATGATCATCTTCTGGACACCTTCCTTCCTGCTCTTCATG      | 60  |
| Db | 257 | ATGAGGGGCAACCCACATGATCATCTTCTGGACATTCCTTCCTCTTCATCTCTCATG    | 316 |
| QY | 61  | GTCGTGCCACGCTGTGCGGACACCCCTGTACCTGTCTTGGACACCCCAAGTGC        | 120 |
| Db | 317 | GTGATATCCCGAGCTGTGGCCAGACACCCCTGTCTCTGTCTTGGACACCCCAAGTGCCA  | 376 |
| QY | 121 | CAGGGGGTACCCCTGTGTGTGGATATGGCTGTGGCTGTCTTAAATGTGTGCAGGAGCTG  | 180 |
| Db | 377 | CCGGGGGTACCCCTGTGTGTGGATATGGCTGTGGCTGTCTTCAAGTGTGCACGGAGGCTG | 436 |
| QY | 181 | GGGAGATCTGTGCAGACACCTGCATGTCTGCACCCACGAGGCGCTGTTTGTAGGCT     | 240 |
| Db | 437 | GGGAGATCTGTGCAGACACCTGCATGTCTGCACCCACGAGGCGCTGTTTGTAGGCT     | 496 |
| QY | 241 | GGGGCAGGCCCCGTGGCGCCATGGGGCTGTGTGTCTTGTGATGAGGATGAGCGTACGT   | 300 |
| Db | 497 | GGGGCAGGCCCCCAGTGGCGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT     | 556 |
| QY | 301 | GAGGTGAATGGCCCGCAGGTACCTGGATGAGAGACCTTTAAACCAATTGCAAGGTCCTG  | 360 |
| Db | 557 | GAGGTGAATGGCCCGCAGGTACCTGGATGAGGAGACCTTTAAACCAATTGCAAGGTTTGG | 616 |
| QY | 361 | TGCCGCTGTATGACGGGTGGCTTACCGGCCCTGGCGCGTGCAGTGAAGATGTGGGGCTG  | 420 |
| Db | 617 | TGCCGCTGTATGACGGGTGGCTTACCTGCTGGCGCGTGTGCAGTGAAGATGTGGGGCTG  | 676 |
| QY | 421 | CCGAGCTGTGGACTCCGACGCCGCCAAGAGAAATACAGTGCAGAGAAATGTGCCCCGAG  | 480 |
| Db | 677 | CCGAGCTGTGGACTCCGACGCCGCCAAGAGAAATACAGTGCAGAGAAATGTGCCCCGAG  | 736 |
| QY | 481 | TGGGTATGTACACGAGGAGTGA---CACCGGATCCAGCGCTCCACGGCGCAAGAGAC    | 537 |
| Db | 737 | TGGGTATGTACACGAGGAGTGAATCAGCGGCAATCCAGCCCTCTTCAGCCCAAGAGAC   | 796 |
| QY | 538 | CAACTTTCGGCCCTGTGCACACGCGCGCTGTATGTCTTGTGCACAAATTTTGGAGACA   | 597 |
| Db | 797 | CAACTTTCGGCCCTGTGTACTCTGTGCATGTCCGATGGCCCTGTGCACAACTGGAGACA  | 856 |
| QY | 598 | GCTGTGGGGCCCTGTCAACACCTGTGTGGCTGTGGGATAGGACCCGAGTGTCAACACAG  | 657 |
| Db | 857 | GCTGTGGGGCCCTGTCTCAACACCTGTGTGGGATAGGACCCGAGTGTCAACACAG      | 916 |
| QY | 658 | AACGATTTGTGCCAATGAGATCAACAGCGCGCTGTGTGTGCCAGACCTGTGCTTGCA    | 717 |
| Db | 917 | AACGATTTGTGCCAATGAGATCAACAGGTGCGCTGTGTGTGTGCCAGACCTGTGCTTGCA | 976 |
| QY | 718 | GCCAGGAGCCACAGCTATGGACAGATGCTTTCTTA                          | 752 |

Db . 977 TCCAGGAGCCACGGCTCATGGAACAGTGCCTCTA 1011

RESULT 5  
US-10-112-267-18/c  
Sequence 18 and 19  
US-10112267

|                           |   |   |
|---------------------------|---|---|
|                           | Publication No.   | US20030068678A1                                   |
|                           | GENERAL INFORMATION:  |   |
|                           | APPLICANT:  | Botstein, David A.                                |
|                           | APPLICANT:  | Cohen, Robert                                     |
|                           | APPLICANT:  | Goddard, Audrey                                   |
|                           | APPLICANT:  | Gurney, Austen L.                                 |
|                           | APPLICANT:  | Hillan, Kenneth J.                                |
|                           | APPLICANT:  | Lawrence, David A.                                |
|                           | APPLICANT:  | Levine, Arnold J.                                 |
|                           | APPLICANT:  | Pennica, Diane                                    |
|                           | APPLICANT:  | Roy, Margaret Ann                                 |
|                           | APPLICANT:  | Wood, William I.                                  |
|                           | TITLE OF INVENTION:   | MISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME |
|                           | FILE REFERENCE:   | PI176R2   |
|                           | CURRENT APPLICATION NUMBER:   | US/10/112,267                                     |
|                           | PRIOR FILING DATE:  | 2002-03-27  |
|                           | PRIOR APPLICATION NUMBER:   | EARLIER APPLICATION NUMBER: US/09/182,145B        |
|                           | PRIOR FILING DATE:  | EARLIER FILING DATE: 1998-10-29                   |
|                           | PRIOR APPLICATION NUMBER:   | EARLIER APPLICATION NUMBER: US 60/063,704         |
|                           | PRIOR FILING DATE:  | EARLIER FILING DATE: 1997-10-29                   |
|                           | PRIOR APPLICATION NUMBER:   | EARLIER APPLICATION NUMBER: US 60/073,612         |
|                           | PRIOR FILING DATE:  | EARLIER FILING DATE: 1998-02-04                   |
|                           | PRIOR APPLICATION NUMBER:   | EARLIER APPLICATION NUMBER: US 60/081,695         |
|                           | PRIOR FILING DATE:  | EARLIER FILING DATE: 1998-04-14                   |
|                           | NUMBER OF SEQ ID NOS:   | 156   |
|                           | SEQ ID NO 18  |   |
|                           | LENGTH:   | 1734  |
|                           | TYPE:   | DNA   |
|                           | ORGANISM:   | Mus musculus                                      |
|                           | US-10-112-267-18  |   |
| Query Match               | 87.5%; Score 659; DB 15; Length 1734;                                   |   |
| Best Local Similarity     | 93.0%; Pred. No. 6.6e-179;  |   |
| Matches 702; Conservative | 0; Mismatches 50; Indels 3; Gaps 1;                                     |   |
| QY                        | 1 ATGAGGGGACGCCACATGCATCTTGTGGCCACTTCCTTCCTCCTTCATAG 60                 |   |
|                           |   |   |
| DB                        | 1478 ATGAGGGGACGCCACATGCATCTTGTGGCCACTTCCTTCCTCCTCATG 1419              |   |
| QY                        | 61 GTGTGTGCCAGTGTCGCCGACACCCTGTACCTGTCTTTGSACACCAACCAGTGCCA 120         |   |
|                           |   |   |
| DB                        | 1418 GTGTATTCCACGCTGTGCCAGCACACCTGTGTCTTGTGAACACCAACCAGTGCCA 1359       |   |
| QY                        | 121 CAGGGGGTACCCCTGGTGTGTGATGCGCTGTGTCTGTAAAGTGTGTGCACGAGGCTG 180       |   |
|                           |   |   |
| DB                        | 1358 CCGGGGGTACCCCTGGTGTGTGATGCGCTGTGTGTGTGTGAGAGTGTGCACGAGGCTG 1299    |   |
| QY                        | 181 GGGGAGTCCCGGACACCACTGCATGTCGTGGACCCCCAGGAGCCGTGTTTGCACGCT 240       |   |
|                           |   |   |
| DB                        | 1298 GGGGAGTCCCTGCACACCACTGCATGTCTGCAGACCCCAGGAGCCGTGTTTGCACGCT 1239    |   |
| QY                        | 241 GGGGACAGCCCTGGCGCCATGGGGCTGTGTCTTCTTGGATGAGATGACGGTAGCTGT 300       |   |
|                           |   |   |
| DB                        | 1238 GGGGACAGCCCTGCAGTGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1179     |   |
| QY                        | 301 GAGGTAAATGGCCGACAGTACCTGTGATGGAGAGAACCTTTAAACCAATTGACAGGTCCTG 360   |   |
|                           |   |   |
| DB                        | 1178 GAGGTAAATGGCCGACAGTACCTGTGATGGAGAGAACCTTTAAACCAATTGACAGGTCCTG 1119 |   |
| QY                        | 361 TGGCGGTGATGACGGTGTGCTTACCTGTGCTGCGCTGTGCAGTGAAGATGTGTGGGCTG 420     |   |
|                           |   |   |
| DB                        | 1118 TGGCGGTGATGACGGTGTGCTTACCTGTGCTGCGCTGTGCAGTGAAGATGTGTGGGCTG 1059   |   |
| QY                        | 421 CCCAGTGGAGCTGCCACGCCCCAAGAGAATAACAGGTGCCAGGAAGTGTGCCCCGAG 480       |   |
|                           |   |   |
| DB                        | 1058 CCCAGTGGAGCTGCCACGCCCCAAGAGAATAACAGGTGCCAGGAAGTGTGCCCCGAG 999      |   |



PRIOR FILING DATE: 1998-03-27  
 PRIOR APPLICATION NUMBER: 60/080165  
 PRIOR FILING DATE: 1998-03-31  
 PRIOR APPLICATION NUMBER: 60/081203  
 PRIOR FILING DATE: 1998-04-09  
 PRIOR APPLICATION NUMBER: 60/081229  
 PRIOR FILING DATE: 1998-04-09  
 PRIOR APPLICATION NUMBER: 60/081695  
 PRIOR FILING DATE: 1998-04-14  
 PRIOR APPLICATION NUMBER: 60/081817  
 PRIOR FILING DATE: 1998-04-15  
 PRIOR APPLICATION NUMBER: 60/081818  
 PRIOR FILING DATE: 1998-04-15  
 PRIOR APPLICATION NUMBER: 60/082999  
 PRIOR FILING DATE: 1998-04-24  
 PRIOR APPLICATION NUMBER: 60/083322  
 PRIOR FILING DATE: 1998-04-28  
 PRIOR APPLICATION NUMBER: 60/083545  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/084600  
 PRIOR FILING DATE: 1998-05-07  
 PRIOR APPLICATION NUMBER: 60/084627  
 PRIOR FILING DATE: 1998-05-07  
 PRIOR APPLICATION NUMBER: 60/084637  
 PRIOR FILING DATE: 1998-05-07  
 PRIOR APPLICATION NUMBER: 60/085149  
 PRIOR FILING DATE: 1998-05-12  
 PRIOR APPLICATION NUMBER: 60/085323  
 PRIOR FILING DATE: 1998-05-13  
 PRIOR APPLICATION NUMBER: 60/085338  
 PRIOR FILING DATE: 1998-05-13  
 PRIOR APPLICATION NUMBER: 60/085339  
 PRIOR FILING DATE: 1998-05-13  
 PRIOR APPLICATION NUMBER: 60/085579  
 PRIOR FILING DATE: 1998-05-15  
 PRIOR APPLICATION NUMBER: 60/085697  
 PRIOR FILING DATE: 1998-05-15  
 PRIOR APPLICATION NUMBER: 60/085704  
 PRIOR FILING DATE: 1998-05-15  
 PRIOR APPLICATION NUMBER: 60/086414  
 PRIOR FILING DATE: 1998-05-22  
 PRIOR APPLICATION NUMBER: 60/086430  
 PRIOR FILING DATE: 1998-05-22  
 PRIOR APPLICATION NUMBER: 60/087106  
 PRIOR FILING DATE: 1998-05-28  
 PRIOR APPLICATION NUMBER: 60/088026  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088730  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088741  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088810  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088858  
 PRIOR FILING DATE: 1998-06-11  
 PRIOR APPLICATION NUMBER: 60/089532  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089599  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089907  
 PRIOR FILING DATE: 1998-06-18  
 PRIOR APPLICATION NUMBER: 60/089947  
 PRIOR FILING DATE: 1998-06-19  
 PRIOR APPLICATION NUMBER: 60/090349  
 PRIOR FILING DATE: 1998-06-23  
 PRIOR APPLICATION NUMBER: 60/090429  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090445  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090538  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090863  
 PRIOR FILING DATE: 1998-06-26

PRIOR APPLICATION NUMBER: 60/091360  
 PRIOR FILING DATE: 1998-07-01  
 PRIOR APPLICATION NUMBER: 60/091519  
 PRIOR FILING DATE: 1998-07-02  
 PRIOR APPLICATION NUMBER: 60/091982

Query Match 67.8%; Score 510.4; DB 14; Length 1266;  
 Best Local Similarity 79.9%; Pred. No. 2,1e-136;  
 Matches 601; Conservative 0; Mismatches 151; Indels 0; Gaps 0;

QY 1 ATGAGGGGACCCACATGATCATCTTGTGGCCACTTCTCTCTCTCTCTCATG 60  
 DB 10 ATGAGGGGACCCACATGATCATCTTGTGGCCACTTCTCTCTCTCTCTCATG 69  
 QY 61 GTGTGTGCCAGCTGTGTGCCAGACCCGTGTGTGTGTGTGTGTGTGTGTGTGT 120  
 DB 70 GTGTGTGCCAGCTGTGTGCCAGACCCGTGTGTGTGTGTGTGTGTGTGTGTGT 129  
 QY 121 CAGGGGGTACCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180  
 DB 130 CTGGGAGTACCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 189  
 QY 181 GGGGAGT 240  
 DB 190 GGGGAGT 249  
 QY 241 GGGGAGT 300  
 DB 250 GGGGAGT 309  
 QY 301 GAGGTGATGT 360  
 DB 310 GAGGTGATGT 369  
 QY 361 TGCCGT 420  
 DB 370 TGCCGT 429  
 QY 421 CCCAGCTGT 480  
 DB 430 CCCAGCTGT 489  
 QY 481 TGGGTATGT 540  
 DB 490 TGGGTATGT 549  
 QY 541 CTTTGT 600  
 DB 550 TTTTGT 609  
 QY 601 TGGGGCCCTGT 660  
 DB 610 TGGGGCCCTGT 669  
 QY 661 CGATTCTGT 720  
 DB 670 CGATTCTGT 729  
 QY 721 AGGAGCCAGCTGT 752  
 DB 730 AGGAGCCAGCTGT 761

RESULT 7  
US-10-146-726-319

Sequence 319, Application US/10146726  
 Publication No. US20030129690A1  
 GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.  
 APPLICANT: Beresini, Maureen  
 APPLICANT: DeForge, Laura  
 APPLICANT: Desnoyers, Luc  
 APPLICANT: Filvaroff, Ellen



```

Db      370  TGCCGCTCGAGAGAGCGGGCTTACCTGCGCGCGCTGTCAGCGAGATGTGCGGTG 429
Qy      421  CCCAGCTGGAGCTGCCCAACGCCCAAGAGATATACAGTGGCCAGAAATGTGCCCCGAG 480
Db      430  CCCAGCTGGAGCTGCCCAACGCCCAAGAGATATACAGTGGCCAGAAATGTGCCCCGAG 489
Qy      481  TGGGTATGTGACCGAGGAGTGCACCGCGGATCCAGCGCTCCACGGCCCAAGACCAAA 540
Db      490  TGGGTGTGCGGCGCAAGAGGGGAGCTGGGAGCCCGCTTCCAGCCCAAGAGACCCAG 549
Qy      541  CTTTGTGCGCTGTGACACCTGCTCTGCTATGCTGTGTCCTCAATTGGAGCACACC 600
Db      550  TTTTGTGCGCTGTGCTCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 609
Qy      601  TGGGCCCCCTGCTCAACCACTGTGGGCTGGGCAATAGCGACCCAGTGTCCAAACAGAC 660
Db      610  TGGGAGCCCTGCTGACACACCTGTGGGCTGGGCAATAGCGACCCAGTGTCCAAACAGAC 669
Qy      661  CGATTCTGCCAAGTGGAGATCCAAACGCCGCTGTGTCTGCTCCAGACCTGCTGCGACCC 720
Db      670  CGCTTCTGCGCACTGGAGACCCAGCGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 729
Qy      721  AGGAGCCACAGCTCATGGAACAGTGCCTTCTA 752
Db      730  AGGGGTGCGAGTCCACAAACAGTGCCTTCTA 761

```

# RESULT 9 US-10-146-788-319

```

; Sequence 319, Application US/10146788
; Publication No. US20030129693A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C322
; CURRENT APPLICATION NUMBER: US/10/146,788
; CURRENT FILING DATE: 2002-05-15
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 319
; LENGTH: 1266
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-146-788-319

```

Query Match 67.8%; Score 510.4; DB 14; Length 1266;  
Best Local Similarity 79.9%; Pred. No. 2,1e-136;  
Matches 601; Conservative 0; Mismatches 151; Indels 0; Gaps 0;

```

Qy      1  ATAGAGGCGACCCACATCATCTTGTGGCCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 60
Db      10  ATGAGAGCGACACCAAGAGCCACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 69
Qy      61  GTGTGTGCGGAGCTGTGCGGAGACCCGTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120

```

```

Db      70  GTGCTACCCACAGCTGTGCCCCGACACACATGTAATCTGCCCCCTGGCCACCTCCCGCATGCCG 129
Qy      121  CAGGAGGTACCCCTGTGTGCTGATGAGTGTGCTGTGTGTAAAGTGTGTGACGAGAGCTG 180
Db      130  CTGGAGAGTACCCCTGTGTGCTGATGAGTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 189
Qy      181  GGGAGTCTTCCGACACCTTCAATGTGTGCAACCCCAAGGAGGAGGCTGTGTGTGTGTGTGTGT 240
Db      190  GGGAGGCTTCCGACACCTTCAATGTGTGCAACCCCAAGGAGGAGGCTGTGTGTGTGTGTGTGT 249
Qy      241  GGGAGGCTTCCGACACCTTCAATGTGTGCAACCCCAAGGAGGAGGCTGTGTGTGTGTGTGTGT 300
Db      250  GGGAGGCTTCCGACACCTTCAATGTGTGCAACCCCAAGGAGGAGGCTGTGTGTGTGTGTGTGT 309
Qy      301  GAGGTGAATGCGCCAGTGTACCTGATGAGAGACCTTTAAACCAATTGACGAGGCTGCTG 360
Db      310  GAGGTGAATGCGCCAGTGTATCGGGAAGAGGAGACCTTCAAGCCCACTGACGATCGCG 369
Qy      361  TGGCGCTGTGATGACGAGTGTGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
Db      370  TGGCGCTGTGAGAGAGCGGCGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 429
Qy      421  CCCAGCTGGAGCTGCCACGCCCAAGAGATACAGGTGCGCAGGAAAGTGTGCGCCGCG 480
Db      430  CCCAGCTGGAGCTGCCACGCCCAAGAGATACAGGTGCGCAGGAAAGTGTGCGCCGCTGAG 489
Qy      481  TGGGTATGTGACCGAGGAGTGCACACCGCGGATCCAGCGCTCCAGCGCGCAAGACCAAA 540
Db      490  TGGGTGTGCGGCGCAAGAGGGGAGCTGGGAGACCCAGCCCTTCCAGCAAGAGACCCAG 549
Qy      541  CTTTGTGCGCTGTGATCATCTGCTGCTGTGTATGCTCTTGTCCAAATGTGAGACAGCC 600
Db      550  TTTTGTGCGCTGTGCTTCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 609
Qy      601  TGGGCCCCCTGCTCAACCACTGTGGGCTGGGCAATAGCGACCCAGTGTCCAAACAGAC 660
Db      610  TGGGAGCCCTGCTGACACACCTGTGGGCTGGGCAATAGCGACCCAGTGTCCAAACAGAC 669
Qy      661  CGATTCTGCCAAGTGGAGATCCAAACGCCGCTGTGTCTGCTCCAGACCTGCTGCGACGCC 720
Db      670  CGCTTCTGCGCACTGGAGACCCAGCGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 729
Qy      721  AGGAGCCACAGCTCATGGAACAGTGCCTTCTA 752
Db      730  AGGGGTGCGAGTCCACAAACAGTGCCTTCTA 761

```

# RESULT 10 US-10-152-380-319

```

; Sequence 319, Application US/10152380
; Publication No. US20030129694A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C397
; CURRENT APPLICATION NUMBER: US/10/152,380

```





Db 610 TGGGGACCTGCTGCGACCTGTGGCTGGGACATGGCCACCCGGGTGTCACACGAGAC 669  
Qy 661 CGATTTCGCACTGAGATCCACAGCCGCTGTGTGTCCAGAGACCTGCGTGGCAGCC 720  
Db 670 CCGTCTGCGGACGTGAGAGACCCAGCCGCTGTGTGTTCACAGGCGCTGCGCTTCTA 729  
Qy 721 AGAGCCACAGCTCATGAGACAGTGTCTTCTA 752  
Db 730 AGGGGTGCGATGTCACAAACAGTGTCTTCTA 761

RESULT 12  
US-10-028-072-319  
Sequence 319, Application US/10028072  
Publication No. US20030004311A1  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: DeForge, Laura  
APPLICANT: Desnoyers, Luc  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang

TITLE OF INVENTION:  
FILE REFERENCE:  
CURRENT APPLICATION NUMBER: US/10/028,072  
CURRENT FILING DATE: 2001-12-19  
PRIOR APPLICATION NUMBER: 60/049911  
PRIOR FILING DATE: 1997-06-18  
PRIOR APPLICATION NUMBER: 60/056974  
PRIOR FILING DATE: 1997-08-26  
PRIOR APPLICATION NUMBER: 60/059113  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/059115  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/059117  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/059122  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/059184  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/059263  
PRIOR FILING DATE: 1997-09-18  
PRIOR APPLICATION NUMBER: 60/059352  
PRIOR FILING DATE: 1997-09-19  
PRIOR APPLICATION NUMBER: 60/059588  
PRIOR FILING DATE: 1997-09-19  
PRIOR APPLICATION NUMBER: 60/059836  
PRIOR FILING DATE: 1997-09-24  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/062285  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/062287  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/062814  
PRIOR FILING DATE: 1997-10-24  
PRIOR APPLICATION NUMBER: 60/062816  
PRIOR FILING DATE: 1997-10-24  
PRIOR APPLICATION NUMBER: 60/063045  
PRIOR FILING DATE: 1997-10-24  
PRIOR APPLICATION NUMBER: 60/063082

PRIOR FILING DATE: 1997-10-31  
PRIOR APPLICATION NUMBER: 60/063127  
PRIOR FILING DATE: 1997-10-24  
PRIOR APPLICATION NUMBER: 60/063327  
PRIOR FILING DATE: 1997-10-27  
PRIOR APPLICATION NUMBER: 60/063329  
PRIOR FILING DATE: 1997-10-27  
PRIOR APPLICATION NUMBER: 60/063550  
PRIOR FILING DATE: 1997-10-28  
PRIOR APPLICATION NUMBER: 60/063561  
PRIOR FILING DATE: 1997-10-28  
PRIOR APPLICATION NUMBER: 60/063704  
PRIOR FILING DATE: 1997-10-29  
PRIOR APPLICATION NUMBER: 60/063733  
PRIOR FILING DATE: 1997-10-29  
PRIOR APPLICATION NUMBER: 60/063735  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/064248  
PRIOR FILING DATE: 1997-11-03  
PRIOR APPLICATION NUMBER: 60/064809  
PRIOR FILING DATE: 1997-11-07  
PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065846  
PRIOR FILING DATE: 1997-11-17  
PRIOR APPLICATION NUMBER: 60/066364  
PRIOR FILING DATE: 1997-11-21  
PRIOR APPLICATION NUMBER: 60/066453  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/066511  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/066770  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/069212  
PRIOR FILING DATE: 1997-12-11  
PRIOR APPLICATION NUMBER: 60/069278  
PRIOR FILING DATE: 1997-12-11  
PRIOR APPLICATION NUMBER: 60/069334  
PRIOR FILING DATE: 1997-12-11  
PRIOR APPLICATION NUMBER: 60/069694  
PRIOR FILING DATE: 1997-12-16  
PRIOR APPLICATION NUMBER: 60/072320  
PRIOR FILING DATE: 1998-01-23  
PRIOR APPLICATION NUMBER: 60/073612  
PRIOR FILING DATE: 1998-02-04  
PRIOR APPLICATION NUMBER: 60/074086  
PRIOR FILING DATE: 1998-02-09  
PRIOR APPLICATION NUMBER: 60/074092  
PRIOR FILING DATE: 1998-02-09  
PRIOR APPLICATION NUMBER: 60/077791  
PRIOR FILING DATE: 1998-03-12  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/079294  
PRIOR FILING DATE: 1998-03-25  
PRIOR APPLICATION NUMBER: 60/079663  
PRIOR FILING DATE: 1998-02-27  
PRIOR APPLICATION NUMBER: 60/079728  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/080165  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/081203  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081229  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081695  
PRIOR FILING DATE: 1998-04-14  
PRIOR APPLICATION NUMBER: 60/081817  
PRIOR FILING DATE: 1998-04-15

|                                       |                               |                                       |                               |                                       |                               |                                       |                               |                                       |                               |                                       |                               |                                       |                               |                                       |                               |                                       |                               |                                       |                               |                                       |                               |                                       |                               |                                       |                               |                                       |                               |                                       |                               |                                       |                               |                                       |                               |                                       |                               |                                       |                               |                                       |                               |                                       |                               |                                       |                               |                                       |                               |                                       |                               |                                       |                               |                                       |                               |                                       |                               |                                       |                               |                                       |                               |                                       |                               |                                       |                               |
|---------------------------------------|-------------------------------|---------------------------------------|-------------------------------|---------------------------------------|-------------------------------|---------------------------------------|-------------------------------|---------------------------------------|-------------------------------|---------------------------------------|-------------------------------|---------------------------------------|-------------------------------|---------------------------------------|-------------------------------|---------------------------------------|-------------------------------|---------------------------------------|-------------------------------|---------------------------------------|-------------------------------|---------------------------------------|-------------------------------|---------------------------------------|-------------------------------|---------------------------------------|-------------------------------|---------------------------------------|-------------------------------|---------------------------------------|-------------------------------|---------------------------------------|-------------------------------|---------------------------------------|-------------------------------|---------------------------------------|-------------------------------|---------------------------------------|-------------------------------|---------------------------------------|-------------------------------|---------------------------------------|-------------------------------|---------------------------------------|-------------------------------|---------------------------------------|-------------------------------|---------------------------------------|-------------------------------|---------------------------------------|-------------------------------|---------------------------------------|-------------------------------|---------------------------------------|-------------------------------|---------------------------------------|-------------------------------|---------------------------------------|-------------------------------|---------------------------------------|-------------------------------|
| PRIOR APPLICATION NUMBER: 60/08181181 | PRIOR FILING DATE: 1998-04-15 | PRIOR APPLICATION NUMBER: 60/08299999 | PRIOR FILING DATE: 1998-04-24 | PRIOR APPLICATION NUMBER: 60/08333222 | PRIOR FILING DATE: 1998-04-28 | PRIOR APPLICATION NUMBER: 60/08354545 | PRIOR FILING DATE: 1998-04-29 | PRIOR APPLICATION NUMBER: 60/08460000 | PRIOR FILING DATE: 1998-05-07 | PRIOR APPLICATION NUMBER: 60/08462727 | PRIOR FILING DATE: 1998-05-07 | PRIOR APPLICATION NUMBER: 60/08465373 | PRIOR FILING DATE: 1998-05-07 | PRIOR APPLICATION NUMBER: 60/08514949 | PRIOR FILING DATE: 1998-05-12 | PRIOR APPLICATION NUMBER: 60/08532323 | PRIOR FILING DATE: 1998-05-13 | PRIOR APPLICATION NUMBER: 60/08533838 | PRIOR FILING DATE: 1998-05-13 | PRIOR APPLICATION NUMBER: 60/08533939 | PRIOR FILING DATE: 1998-05-13 | PRIOR APPLICATION NUMBER: 60/08557979 | PRIOR FILING DATE: 1998-05-15 | PRIOR APPLICATION NUMBER: 60/08569797 | PRIOR FILING DATE: 1998-05-15 | PRIOR APPLICATION NUMBER: 60/08570404 | PRIOR FILING DATE: 1998-05-15 | PRIOR APPLICATION NUMBER: 60/08641414 | PRIOR FILING DATE: 1998-05-22 | PRIOR APPLICATION NUMBER: 60/08643030 | PRIOR FILING DATE: 1998-05-22 | PRIOR APPLICATION NUMBER: 60/08710606 | PRIOR FILING DATE: 1998-05-28 | PRIOR APPLICATION NUMBER: 60/08802626 | PRIOR FILING DATE: 1998-06-04 | PRIOR APPLICATION NUMBER: 60/08887300 | PRIOR FILING DATE: 1998-06-10 | PRIOR APPLICATION NUMBER: 60/08874111 | PRIOR FILING DATE: 1998-06-10 | PRIOR APPLICATION NUMBER: 60/08884040 | PRIOR FILING DATE: 1998-06-10 | PRIOR APPLICATION NUMBER: 60/08885858 | PRIOR FILING DATE: 1998-06-11 | PRIOR APPLICATION NUMBER: 60/08935222 | PRIOR FILING DATE: 1998-06-17 | PRIOR APPLICATION NUMBER: 60/08959999 | PRIOR FILING DATE: 1998-06-17 | PRIOR APPLICATION NUMBER: 60/09042929 | PRIOR FILING DATE: 1998-06-24 | PRIOR APPLICATION NUMBER: 60/09044545 | PRIOR FILING DATE: 1998-06-24 | PRIOR APPLICATION NUMBER: 60/09053838 | PRIOR FILING DATE: 1998-06-24 | PRIOR APPLICATION NUMBER: 60/09086333 | PRIOR FILING DATE: 1998-06-26 | PRIOR APPLICATION NUMBER: 60/09131606 | PRIOR FILING DATE: 1998-07-01 | PRIOR APPLICATION NUMBER: 60/09151919 | PRIOR FILING DATE: 1998-07-02 | PRIOR APPLICATION NUMBER: 60/09196222 | PRIOR FILING DATE: 1998-07-07 |
|---------------------------------------|-------------------------------|---------------------------------------|-------------------------------|---------------------------------------|-------------------------------|---------------------------------------|-------------------------------|---------------------------------------|-------------------------------|---------------------------------------|-------------------------------|---------------------------------------|-------------------------------|---------------------------------------|-------------------------------|---------------------------------------|-------------------------------|---------------------------------------|-------------------------------|---------------------------------------|-------------------------------|---------------------------------------|-------------------------------|---------------------------------------|-------------------------------|---------------------------------------|-------------------------------|---------------------------------------|-------------------------------|---------------------------------------|-------------------------------|---------------------------------------|-------------------------------|---------------------------------------|-------------------------------|---------------------------------------|-------------------------------|---------------------------------------|-------------------------------|---------------------------------------|-------------------------------|---------------------------------------|-------------------------------|---------------------------------------|-------------------------------|---------------------------------------|-------------------------------|---------------------------------------|-------------------------------|---------------------------------------|-------------------------------|---------------------------------------|-------------------------------|---------------------------------------|-------------------------------|---------------------------------------|-------------------------------|---------------------------------------|-------------------------------|---------------------------------------|-------------------------------|

|                           |        |                     |           |              |
|---------------------------|--------|---------------------|-----------|--------------|
| Query Match               | 67.8%; | Score 510.4;        | DB 15;    | Length 1266; |
| Best Local Similarity     | 79.9%; | Pred. No. 2.1e-136; |           |              |
| Matches 601; Conservative | 0;     | Mismatches 151;     | Indels 0; | Gaps 0;      |

|   |     |  |     |
|---|-----|--|-----|
| QY                                      | 1   | ATGAGGGGAGGAGCCACTGATGATCCATCTCTGGGCACTTCCCTCTCTGCTGCTGCTCATG    | 60  |
| Db                                      | 10  | ATGAGGAGGACACCGAAGACCCACTCTCTGGGCTTCTCTCCCTCTCTGCTCTCTCTCAAG     | 69  |
| QY                                      | 61  | GTGTGTGGCCACTGTGTCGGAGACACCTGTACTGTCTGCTTGGAGACCAACCCAGTGC       | 120 |
| Db                                      | 70  | GTGCGTACCCAGCTGTGCCCCGACACCATGTACTGTCCCTGTGGCACCCTCCCCGATGGCCG   | 129 |
| QY                                      | 121 | CAGGGGGTACCCCTGGTGTGATGGCTGTGGCTGTCTTAAAGTGTGTGCAGCGAGGCTGT      | 180 |
| Db                                      | 130 | CTGGGAGTACCCCTGGTGTGATGGCTGTGGCTGTCTCCGGGTATGTGACGGGAGCTGT       | 189 |
| QY                                      | 181 | GGGGAGTCTCTGGAGACCACTGATGTCTGTGCCACCCACCCAGGGGCTGTGTGTACACT      | 240 |
| Db                                      | 190 | GGGGAGCCCTGTGGACCACTCCACTGTGTGTGCGACGGCAGACCCAGGGGCTGTGTGTGCA    | 249 |
| QY                                      | 241 | GGGGCAGGGCCCTGGCGGGCCATGGGGCTGTGTGTCTCTTGGATGTAGAGATGAGTGTGT     | 300 |
| Db                                      | 250 | GGGGGACGAGACCCGGGTGTGGCCGGGGGGCCCTGTGTCCCTTGTGGACAGAGACAGCACTGT  | 309 |
| QY                                      | 301 | GAGGTGATGGCGCGCAGGTACTGTGATGTGAGAGACTTTTAAACCCCAATTGCAGGGGTCTG   | 360 |
| Db                                      | 310 | GAGGTGAACGGCGCCCTGTATGGGGAAGGGGAGACTTTCAGACCCCACTGCACATCTCCG     | 369 |
| QY                                      | 361 | TGCGCGTGTGATGACGTGTGGCTTTCACATCTGCTGCCGCTGTGTGACGTGAGGATGTGGCTGT | 420 |
| Db                                      | 370 | TGCGCGTGTGAGAGAGCGGCTTTCACCTGTGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGT    | 429 |
| QY                                      | 421 | CCGAGCTGGGAGCTGCCACGCCCCCAAGAGAAATACAGGTGTGGCAGGAAAGTGTGTGCCGAG  | 480 |
| Db                                      | 430 | CCGAGCTGGGAGCTGCCACCCCAAGAGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT    | 489 |
| QY                                      | 481 | TGGGTATGTGACAGGGAGTGTACACGGCGATCCAGCGCTTCACAGGGCGCAGAGACCA       | 540 |
| Db                                      | 490 | TGGGTGTGGCGCCAGAGAGGGGAGCTGGGGAGACCGACCCCTTCCAGCCCAAGGAGCCAG     | 549 |
| QY                                      | 541 | CTTTCCTGCCCTGTCTACTCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT      | 600 |
| Db                                      | 550 | TTTTTCTGGCGCTGTCT    | 609 |
| QY                                      | 601 | TGGGGGCGCCCTGGCTCAACCACTGTGTGGCTGGGCGATAGCCACCGAGTGTCCACAGAAC    | 660 |
| Db                                      | 610 | TGGGGAGCCCTGCTTCAGACCACTGTGTGGCTGTGGCGATGTGGCACCCGGGTGTCCACCA    | 669 |
| QY                                      | 661 | CGATTTGTGCAACTGTGAGATCCACGCGCCTGTGTCTGTGCCAGAACCTGTGCTGGACCC     | 720 |
| Db                                      | 670 | CGCTTGTGCGACTGTGAGACCCAGGCGCGCTGTGTCTGTGTGTGTGTGTGTGTGTGTGT      | 729 |
| QY                                      | 721 | AGGAGCCACAGCTCATGTGACACAGTCTTTCTA                                | 752 |
| Db                                      | 730 | AGGGGTGTGCAATCCAAACAGTCTTTCTA                                    | 761 |
| RESULT 13                               |     |  |     |
| US-10-121-049-319                       |     |  |     |
| ; Sequence 319, Application US/10121049 |     |  |     |
| ; Publication No. US20030022239A1       |     |  |     |
| GENERAL INFORMATION:                    |     |  |     |
| APPLICANT: Baker, Kevin P.              |     |  |     |
| APPLICANT: Beresini, Maureen            |     |  |     |
| APPLICANT: DeForge, Laura               |     |  |     |
| APPLICANT: Desnoyers, Luc               |     |  |     |
| APPLICANT: Filvaroff, Ellen             |     |  |     |
| APPLICANT: Gao, Wei-Qiang               |     |  |     |
| APPLICANT: Gerltzen, Mary E.            |     |  |     |
| APPLICANT: Goddard, Audrey              |     |  |     |
| APPLICANT: Godowski, Paul J.            |     |  |     |
| APPLICANT: Gurney, Austin L.            |     |  |     |
| APPLICANT: Sherwood, Steven             |     |  |     |
| APPLICANT: Smith, Victoria              |     |  |     |
| APPLICANT: Stewart, Timothy A.          |     |  |     |
| APPLICANT: Tumas, Daniel                |     |  |     |
| APPLICANT: Watanabe, Colin K            |     |  |     |

```

; APPLICANT: Wood,William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C17
; CURRENT APPLICATION NUMBER: US/10/121,049
; PRIORITY FILING DATE: 2002-04-12
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 319
; LENGTH: 1266
; TYPE: DNA
; ORGANISM: Homo Sapien
; US-10-121-049-319

```

```

Query Match      67.8%; Score 510.4; DB 15; Length 1266;
Best Local Similarity 79.9%; Pred. No. 2,1e-136;
Matches 601; Conservative 0; Mismatches 151; Indels 0; Gaps 0;

```

```

QY 1 ATGAGGGGAGCCCACTGATCCATCTTGGCCACTTGGTCTGCTGCTCTCAATG 60
DB 10 ATGAGAGGACACCGAAGACCCACTCGGCTTCTGCTGCTGCTGCTGCTCAAG 69
QY 61 GTGTGTGCCACTGTGCCGAGACCCCTGTACTCTGCTGAGACACACCCAGTCCCA 120
DB 70 GTGCTTACCCACTGTGCGGACACCATGTACTCTGCTGCGACCTCCGATGCCCG 129
QY 121 CAGGGGTACCCCTGTGCTGTGATGGCTGTGCTGCTGCTGCTGCTGCTGCTG 180
DB 130 CTGGAGATACCCCTGTGCTGTGATGGCTGTGCTGCTGCTGCTGCTGCTGCTG 189
QY 181 GGGAGTCTGTGCGACACCATGTATGTGCGACCCCAAGGAGGAGGAGGAGGAG 240
DB 190 GGGGAGCCCTGTGCGACCACTGTGCGACCCCAAGGAGGAGGAGGAGGAGGAG 249
QY 241 GGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 300
DB 250 GGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 309
QY 301 GAGGTGAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360
DB 310 GAGGTGAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 369
QY 361 GTGTGTGCCACTGTGCCGAGACCCCTGTACTCTGCTGAGACACCCAGTCCCA 420
DB 370 GTGCTTACCCACTGTGCGGACACCATGTACTCTGCTGCGACCTCCGATGCCCG 429
QY 421 CCCAGTGGAGTGGCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
DB 430 CCCAGTGGAGTGGCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 489
QY 481 TGGGTATGTGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540
DB 490 TGGGTATGTGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 549
QY 541 CTTTGTGCCCTGTGTACTCTGCTGTGATGCTCTGCTGCTGCTGCTGCTGCTG 600
DB 550 TTTTGTGCCCTGTGTACTCTGCTGTGATGCTCTGCTGCTGCTGCTGCTGCTG 609
QY 601 TGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
DB 610 TGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 669
QY 661 CGATTCTGCAACTGAGATCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720
DB 670 CGATTCTGCAACTGAGATCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 729
QY 721 AGGAGCCACAGCTCATGGAACAGTCTTTCTA 752
DB 730 AGGAGCCACAGCTCATGGAACAGTCTTTCTA 761

```

RESULT 14

```

US-10-123-904-319
; Sequence 319, Application US/10123904
; Publication No. US20030022328A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gueney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C54
; CURRENT APPLICATION NUMBER: US/10/123,904
; PRIORITY FILING DATE: 2002-04-16
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 319
; LENGTH: 1266
; TYPE: DNA
; ORGANISM: Homo Sapien
; US-10-123-904-319

```

```

Query Match      67.8%; Score 510.4; DB 15; Length 1266;
Best Local Similarity 79.9%; Pred. No. 2,1e-136;
Matches 601; Conservative 0; Mismatches 151; Indels 0; Gaps 0;

```

```

QY 1 ATGAGGGGAGCCCACTGATCCATCTTGGCCACTTGGTCTGCTGCTCTCAATG 60
DB 10 ATGAGAGGACACCGAAGACCCACTCGGCTTCTGCTGCTGCTGCTGCTCAAG 69
QY 61 GTGTGTGCCACTGTGCCGAGACCCCTGTACTCTGCTGAGACACCCAGTCCCA 120
DB 70 GTGCTTACCCACTGTGCGGACACCATGTACTCTGCTGCGACCTCCGATGCCCG 129
QY 121 CAGGGGTACCCCTGTGCTGTGATGGCTGTGCTGCTGCTGCTGCTGCTGCTG 180
DB 130 CTGGAGATACCCCTGTGCTGTGATGGCTGTGCTGCTGCTGCTGCTGCTGCTG 189
QY 181 GGGAGTCTGTGCGACACCATGTATGTGCGACCCCAAGGAGGAGGAGGAGGAG 240
DB 190 GGGGAGCCCTGTGCGACCACTGTGCGACCCCAAGGAGGAGGAGGAGGAGGAG 249
QY 241 GGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 300
DB 250 GGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 309
QY 301 GAGGTGAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360
DB 310 GAGGTGAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 369
QY 361 TGGCGCTGTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420
DB 370 TGGCGCTGTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 429
QY 421 CCCAGTGGAGTGGCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
DB 430 CCCAGTGGAGTGGCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 489
QY 481 TGGGTATGTGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540
DB 490 TGGGTATGTGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 549

```





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 18:02:17 ; Search time 1335.24 Seconds  
(without alignments)  
9133.322 Million cell updates/sec

Title: US-10-010-408-3  
Perfect score: 753  
Sequence: 1 ATGAGGGGACAGCCACTGAT.....CATGAACAGCTTCTTAA 753

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues  
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: em\_estb:.\*  
2: em\_esthum:.\*  
3: em\_estlin:.\*  
4: em\_estlinu:.\*  
5: em\_estroy:.\*  
6: em\_estcpl:.\*  
7: em\_estro:.\*  
8: em\_hlc:.\*  
9: gb\_estl:.\*  
10: gb\_estl2:.\*  
11: gb\_hlc:.\*  
12: gb\_estl3:.\*  
13: gb\_estl4:.\*  
14: gb\_estl5:.\*  
15: em\_estfun:.\*  
16: em\_estom:.\*  
17: gb\_gss:.\*  
18: em\_gss\_hum:.\*  
19: em\_gss\_inv:.\*  
20: em\_gss\_pln:.\*  
21: em\_gss\_vrt:.\*  
22: em\_gss\_fun:.\*  
23: em\_gss\_mam:.\*  
24: em\_gss\_mus:.\*  
25: em\_gss\_other:.\*  
26: em\_gss\_pro:.\*  
27: em\_gss\_rnd:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 1          | 508.8 | 67.6        | 1058   | 14 | BM805088    |
| 2          | 500   | 66.4        | 979    | 14 | BO279131    |
| 3          | 463.6 | 61.6        | 886    | 13 | BI822142    |
| 4          | 430.4 | 57.2        | 750    | 13 | BM043988    |
| 5          | 430.2 | 57.1        | 651    | 13 | BI457141    |
| 6          | 427   | 56.7        | 916    | 13 | BI457367    |

|    |       |      |      |    |          |          |           |
|----|-------|------|------|----|----------|----------|-----------|
| 7  | 421.2 | 55.9 | 1006 | 14 | BM921531 | BM921531 | AGENCOURT |
| 8  | 417.8 | 55.5 | 1073 | 14 | BO073722 | BO073722 | AGENCOURT |
| 9  | 409   | 54.3 | 800  | 13 | BI826781 | BI826781 | 603077268 |
| 10 | 404.2 | 53.7 | 928  | 13 | BI161474 | BI161474 | 602864871 |
| 11 | 398.8 | 53.0 | 888  | 13 | BI825652 | BI825652 | 603072631 |
| 12 | 387.2 | 51.4 | 790  | 13 | BM046275 | BM046275 | 603626068 |
| 13 | 375.2 | 49.8 | 1022 | 14 | BO952960 | BO952960 | AGENCOURT |
| 14 | 365.4 | 48.5 | 620  | 13 | BG928868 | BG928868 | HNC57-1-D |
| 15 | 332   | 44.1 | 1251 | 14 | BO961357 | BO961357 | AGENCOURT |
| 16 | 330.8 | 43.9 | 1166 | 13 | BM543799 | BM543799 | AGENCOURT |
| 17 | 320.8 | 42.6 | 489  | 14 | BM751866 | BM751866 | K-EST0028 |
| 18 | 319.2 | 42.4 | 749  | 9  | AL555144 | AL555144 |           |
| 19 | 296.8 | 39.4 | 933  | 14 | BO278961 | BO278961 | AGENCOURT |
| 20 | 293.2 | 38.9 | 452  | 10 | BE481184 | BE481184 | 166529 BA |
| 21 | 251   | 33.3 | 940  | 14 | BO937887 | BO937887 | AGENCOURT |
| 22 | 234.4 | 31.1 | 657  | 13 | BM488489 | BM488489 | P9M20.PK0 |
| 23 | 223.2 | 29.6 | 405  | 12 | BG900069 | BG900069 | H0A51-1-A |
| 24 | 222   | 29.5 | 618  | 12 | BG538655 | BG538655 | 602566932 |
| 25 | 210.4 | 27.9 | 537  | 14 | BO560868 | BO560868 | H4067A01- |
| 26 | 209.6 | 27.8 | 380  | 12 | BG900020 | BG900020 | H0A48-1-G |
| 27 | 193.8 | 25.7 | 424  | 10 | BB849097 | BB849097 |           |
| 28 | 193.2 | 25.7 | 792  | 13 | BI823588 | BI823588 | 603040962 |
| 29 | 182.6 | 25.6 | 401  | 12 | BF849358 | BF849358 | CM3-EN007 |
| 30 | 171.6 | 22.8 | 668  | 9  | AL555143 | AL555143 |           |
| 31 | 167.4 | 22.2 | 966  | 12 | BF141695 | BF141695 | 601790752 |
| 32 | 162.8 | 21.6 | 823  | 14 | BO067127 | BO067127 | AGENCOURT |
| 33 | 154.4 | 20.5 | 948  | 9  | AL542806 | AL542806 |           |
| 34 | 153.4 | 20.4 | 879  | 9  | AL547439 | AL547439 |           |
| 35 | 142   | 18.9 | 1076 | 13 | BM554624 | BM554624 | AGENCOURT |
| 36 | 135.4 | 18.0 | 1002 | 9  | AL556435 | AL556435 |           |
| 37 | 132.8 | 17.6 | 896  | 14 | BO721003 | BO721003 | AGENCOURT |
| 38 | 128.2 | 17.0 | 1117 | 12 | BF982720 | BF982720 | 602304922 |
| 39 | 127.8 | 17.0 | 936  | 14 | BO880159 | BO880159 | AGENCOURT |
| 40 | 127.4 | 16.9 | 887  | 14 | BO894536 | BO894536 | AGENCOURT |
| 41 | 127   | 16.9 | 887  | 12 | BE898435 | BE898435 | 601681295 |
| 42 | 126.8 | 16.8 | 1085 | 13 | BM558306 | BM558306 | AGENCOURT |
| 43 | 126.6 | 16.8 | 828  | 9  | AL543019 | AL543019 |           |
| 44 | 125.2 | 16.6 | 657  | 9  | AL711224 | AL711224 | DKF2P686B |
| 45 | 125   | 16.6 | 505  | 12 | BF037009 | BF037009 | 601456766 |

ALIGNMENTS

RESULT 1  
BM805088  
LOCUS  
DEFINITION  
AGENCOURT\_6490429 NIH\_MGC\_125 Homo sapiens cDNA clone IMAGE:5725937  
5', mRNA sequence.  
ACCESSION  
BM805088  
VERSION  
BM805088.1 GI:19121911  
KEYWORDS  
EST.  
SOURCE  
human.  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 (bases 1 to 1058)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
JOURNAL  
Contact: Robert Strausberg, Ph.D.  
COMMENT  
Email: gcgabs-remail.nih.gov  
Tissue Procurement: Invitrogen  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNLN at:  
<http://image.llnl.gov>  
Plate: LNAK12717 row: j column: 18  
High quality sequence stop: 697.  
Location/Qualifiers  
1..1058

FEATURES  
source

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5725937"
/lab_host="NIH_MGC_125"
/Note="Organ: ovary (pool of 3); Vector: pCMV-Sport6;
Site_1: EcoRV (destroyed); Site_2: NotI; RNA source pool
of three ovaries, from females ranging in age from 38 to
49 yo. Library is oligo-dT primed and directionally cloned
(EcoRV site is destroyed upon cloning). Average insert
size 2.1 kb. Insert size range 1-3.5 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 036."
BASE COUNT      139 a      423 c      294 g      200 t      2 others
ORIGIN

```

```

Query Match      67.6%; Score 508.8; DB 14; Length 1058;
Best Local Similarity 79.8%; Pred. No. 1.2e-114;
Matches 600; Conservative 0; Mismatches 152; Indels 0; Gaps 0;

```

```

1 ATGAGGGGACCCCACTGATCATCTTGTGGCACTTCTGCTCTCTCATG 60
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
9 ATGAGAGGACACCGAAGACCACCTCTGCTCTCTCTCTCTCTCTCAAG 68
61 GTGTGTGCGCAGCTGTGCGGACACCCCTGACCTGTCTTGGACACCCAGTCCCA 120
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
69 GTGGGTACCCAGCTGTGCGGACACCATGTACCTGCCCCGCGCCAGATGCCG 128
121 CAGGGGGTACCCCTGTGCTGTGATGCTGTGCTGTGCTGTGCTGTGCTGTG 180
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
129 CTGGAGTACCCCTGTGCTGTGATGCTGTGCTGTGCTGTGCTGTGCTGTG 188
181 GGGAGTCTCTGACACCATCTGATGTGACACCCAGGAGGCGCTGTGCTGTG 240
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
189 GGGAGGCTCTGACACCATCTGATGTGACACCCAGGAGGCGCTGTGCTGTG 248
241 GGGGAGGCGCTGTGCGGACATGGGCTGTGCTGTGCTGTGCTGTGCTGTG 300
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
249 GGGGAGGAGCCCGGTGAGAGGGGGCGCTGTGCTGTGCTGTGCTGTGCTGT 308
301 GAGGTGATGAGCCCGACGTACTGTGATGAGAGACCTTAAACCAATTGACAGGT 360
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
309 GAGGTGAAAGGCGCGCTGTATGCGAAGGAGGAGACCTTGAAGCCCATGACAT 368
361 TGCCGCTGTGATGAGAGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 420
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
369 TGCCGCTGTGAGAGAGGCGGCTTCACTGTGCTGTGCTGTGCTGTGCTGTG 428
429 CCCAGCTGTGAGAGGCGGCGGACCCAGAGAGGCTGTGAGTGTGCTGTGCTGTG 488
481 TGCGTATGTGACAGAGGAGTACACCGGCTGTGCTGTGCTGTGCTGTGCTGTG 540
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
489 TGGGTGTGCGGCGCAAGAGGAGGAGGAGTGTGCTGTGCTGTGCTGTGCTGTG 548
541 CTCTTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 600
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
549 TTTTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 608
601 TGGGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 660
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
609 TGGGAGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 668
661 CGATTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 720
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
669 CGCTTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 728
721 AGGAGCCACAGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 752
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
729 AGGGGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 760

```

```

RESULT 2
BQ279131
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

```

BQ279131 979 bp mRNA linear EST 07-MAY-2002  
 AGENCOURT 7046721 NIH\_MGC\_107 Homo sapiens cDNA clone IMAGE:5805819  
 5' mRNA sequence.  
 BQ279131  
 BQ279131.1 GI:20489339  
 EST.  
 SOURCE  
 human.  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NIH-MGC <http://imgc.ncl.nih.gov/>.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Rubin Laboratory  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LLCM2045 row: k column: 04  
 High quality sequence stop: 752.

## FEATURES

source  
 location/qualifiers  
 1..979  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5805819"  
 /clone\_lib="NIH\_MGC\_107"  
 /tissue\_type="adenocarcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /Note="Organ: breast; Vector: pOT7; Site\_1: EcoRI;  
 Site\_2: XhoI; CDNA made by oligo-dT priming.  
 Directionally cloned into EcoRI/XhoI sites using the  
 following 5' adaptor: GGCACAG(G). Library constructed by  
 Ling Hong in the laboratory of Gerald M. Rubin (University  
 of California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies).  
 Note: this is a NIH\_MGC library."

```

BASE COUNT      141 a      347 c      313 g      170 t      8 others
ORIGIN

```

```

Query Match      66.4%; Score 500; DB 14; Length 979;
Best Local Similarity 80.2%; Pred. No. 1.6e-112;
Matches 587; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

```

```

21 CCATCTTGTGCGCCTTCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 80
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
13 CCACCTCTGTGCGCTTCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 72
81 GACACCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 140
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
73 GACACATGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 132
141 GGATGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 200
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
133 GGATGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 192
201 GCATGTGTGCGACCCAGGCGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 260
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
193 CCAGCTGTGCGACCGCAGGCGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 252
261 TGGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 320
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
253 GGGGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 312
321 CTTGATGTGAGAGACCTTAAACCAATTGACAGGCTGTGCTGTGCTGTGCTGTGCTGTG 380
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```



```

Db      313 TCGGAGAGGAGACACCTCCAGCCCACTGACATCCGCTGCCGTGCGAGGACGGCG 372
QY      381 CTTCACCTGCGCTGCGCTGTCAGTAGAGATGCGCGCTGCCAGCTGGAGTGGCCACG 440
Db      373 CTTCACCTGCGCTGCGCTGTCAGTAGAGATGCGCGCTGCCAGCTGGAGTGGCCACG 432
QY      441 CCCCAAGAGATACAGGTGTCAGAGAAAGTCTGCCGATGATGTCAGTACAGGAGT 500
Db      433 CCCCAAGAGGTGAGAGTCTGTCAGAGAGTCTGCCGATGATGTCAGTACAGGAGT 492
QY      501 GACACCGGCGATCCAGCGCTCCAGCGGCGCAGAGACCAACTTTCGCCCTGTCACTCC 560
Db      493 GGCAGTGGGAGACCCCTTCACAGGAGAGAGTTCCTGAGCTTCTCTCTTC 552
QY      561 TGCCTGCTGATGCTCTTGTTCAGATGAGACAGCTGGGCGCCCTGCTCAACAC 620
Db      553 CCGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 612
QY      621 CTGTGGGCTGGGATAGCCACCCGAGTGTCCAAACCAACGATTCGTGCACTGAGAT 680
Db      613 CTGTGGGCTGGGATAGCCACCCGAGTGTCCAAACCAACGATTCGTGCACTGAGAT 672
QY      681 CCAGCCCGGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 740
Db      673 CCAGCCCGGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 732
QY      741 CAGTGTCTCTA 752
Db      733 CAGTGTCTCTA 744

```

```

RESULT 3
BI822142 886 bp mRNA linear EST 04-OCT-2001
LOCUS 603039845F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5180650 5'
DEFINITION
BI822142
mRNA sequence.
BI822142.1 GI:15933692
EST.
SOURCE human.
ORGANISM Homo sapiens

```

```

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE NIH-MGC hit://mgc.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM1450 row: j column: 11
High quality sequence stop: 865.
Location/Qualifiers
1. 886

```

## FEATURES

```

source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5180650"
/clone_lib="NIH_MGC_115"
/lab_host="DH10B"
/Note="Organ: pooled brain, lung, testis; Vector:
pcmv-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.

```

```

BASE COUNT 134 a 310 c 288 g 154 t
ORIGIN

```

```

Query Match 61.6%; Score 463.6; DB 13; Length 886;
Best Local Similarity 77.2%; Pred. No. 1.4e-103;
Matches 576; Conservative 0; Mismatches 169; Indels 1; Gaps 1;

```

```

QY      7 GGCAGCCCACTGATTCATCTTCTGCGCACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 66
Db      1 GGCAGCCCACTGATTCATCTTCTGCGCACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 60
QY      67 GGCAGCCCACTGATTCATCTTCTGCGCACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 126
Db      61 ACCAGCTGTGCTGCGCACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120
QY      127 GTACCCCTGATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 186
Db      121 GTACCCCTGATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
QY      187 TCTGCGACCACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 246
Db      181 CCTGCGACCACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
QY      247 GGCCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 306
Db      241 GACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
QY      307 AATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 366
Db      301 AATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
QY      367 TGTGATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 426
Db      361 TGTGATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
QY      427 TGTGATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 486
Db      421 TGTGATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
QY      487 TGTGATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 546
Db      481 TGTGATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 539
QY      547 GGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 606
Db      540 GGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 599
QY      607 CCTGCTCAACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 666
Db      600 CCTGCTCAACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 659
QY      667 TGTGATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 726
Db      660 TGTGATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 719
QY      727 CAGAGCTCATGAGACAGTCTTCTA 752
Db      720 CAGAGCTCATGAGACAGTCTTCTA 745

```

```

RESULT 4
LOCUS BM043988 750 bp mRNA linear EST 07-NOV-2001
DEFINITION 603620978F1 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:5446794 5'
ACCESSION BM043988
VERSION BM043988.1 GI:16773255
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```





SOURCE  
ORGANISM human.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
AUTHORS NIH-MGC http://mgc.ncl.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT  
Contact: Robert Strausberg, Ph.D.  
Email: cga@bbs-remail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.lnl.gov  
Plate: LLM12788 row: b column: 18  
High quality sequence stop: 593.  
Location/Qualifiers  
1. 1006  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5753009"  
/clone\_lib="NIH-MGC\_115"  
/lab\_host="DH10B"  
/note="Organ: pooled brain, lung, testis; Vector:  
pCMV-SPORT6; Site:1: NotI; Site:2: EcoRI (destroyed); RNA  
source anonymous pool of 6 male brains, age range 23-27; 1  
male lung, age 27; and 1 male testis, age 65. Library is  
oligo-dT primed and directionally cloned (EcoRI site is  
destroyed upon cloning). Average insert size 1.8 kb,  
insert size range 1-3 kb. Library is normalized and  
enriched for full-length clones and was constructed by C.  
Gruber (Invitrogen). Research Genetics tracking code  
021. Note: this is a NIH MGC Library."

BASE COUNT  
ORIGIN 148 a 368 c 317 g 109 t 4 others

Query Match 55.5%; Score 421.2; DB 14; Length 1006;  
Best Local Similarity 77.5%; Pred. No. 3.9e-93;  
Matches 585; Conservative 0; Mismatches 161; Indels 9; Gaps 6;

8 GCACCCAGATGATCTTCTGCGCCTCTCTGCTCTCTCAATGCTGTG 67  
1 GCACACGAGAACCCACTCTCTGCTCTCTCTCTCTCTCAAGGTCGTA 60  
68 CCAGCTGTGCGGACACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 127  
61 CCCAGCTGTGCGGACACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120  
128 TACCCCTGTGCTGTGATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 187  
121 TACCCCTGTGCTGTGATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 180  
188 CTTGCGACACCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 247  
181 CTTGCGACACCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 240  
248 GCGCTGTGCGGACACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 307  
241 GACCCGGTGTGCGGACACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300  
308 ATGGCCGAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 367  
301 ACGGCCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 360  
368 GTATGACGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 427  
361 GCGAGGAGCGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 420  
428 GGGACTGTGCGGACACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 487

Db 421 GGGACTGTGCGGACACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 480  
Qy 488 GTGACCAAGAGATGATACACCGGCGATTCACCGGCGCAAGACCACTTCTG 547  
Db 481 GCGGCGAAGAGAGGAGTGTGCGGACCGGCGGCTTCCAGCCCAAGACCACTTCTG 540  
Qy 548 CCGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 606  
Db 541 GCGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 600  
Qy 607 CCGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 661  
Db 601 CCGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 660  
Qy 662 GATTCGCAACAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 720  
Db 661 GCTTGTGCGGACGATGATGATGATGATGATGATGATGATGATGATGATGAT 720  
Qy 721 A-GGAGCCACAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 753  
Db 721 AGGGGGGTGCGGACGATGATGATGATGATGATGATGATGATGATGATGAT 755

RESULT 8  
B0073722 1073 bp mRNA linear EST 02-APR-2002  
LOCUS B0073722  
DEFINITION AGENCOURT\_7046577 NIH\_MGC\_101 Homo sapiens CDNA clone IMAGE:5806602  
ACCESSION B0073722  
VERSION B0073722.1 GI:19902768  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
AUTHORS NIH-MGC http://mgc.ncl.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT  
Contact: Robert Strausberg, Ph.D.  
Email: cga@bbs-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.lnl.gov  
Plate: LLM2047 row: k column: 19  
High quality sequence stop: 517.  
Location/Qualifiers  
1. 1073  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5806602"  
/clone\_lib="NIH-MGC\_101"  
/tissue\_type="epidermoid carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: lung; Vector: pORF7; Site:1: EcoRI; Site:2:  
XhoI; CDNA made by oligo-dT priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GGCAGAGAG(6). Library constructed by Ling Hong in the  
laboratory of Gerald M. Rubin (University of California,  
Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and  
Superscript II RT (Life Technologies). Note: this is a  
NIH-MGC Library."

BASE COUNT  
ORIGIN 166 a 384 c 347 g 176 t

Query Match 55.5%; Score 417.8; DB 14; Length 1073;  
Best Local Similarity 79.5%; Pred. No. 2.8e-92;  
Matches 532; Conservative 0; Mismatches 132; Indels 5; Gaps 3;



|                       |  |   |                             |
|-----------------------|--|---|-----------------------------|
| Db                    | 668  | CNAGCGGCGCTGTGCCTTCACAGGCCCTTGCCACCCTCAGGGGTCGACAGTCCACAACAA      | 727                         |
| RESULT 10             | Bt161474   | 928 bp  | mRNA linear EST 05-JUL-2001 |
| LOCUS                 | Bt161474   | 60286487f1 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:5019065 5'    |                             |
| DEFINITION            | Bt161474   | mRNA sequence.  |                             |
| ACCESSION             | Bt161474   | GI:14621475   |                             |
| VERSION               | Bt161474.1   | GI:14621475   |                             |
| KEYWORDS              | EST.   |   |                             |
| SOURCE                | human.   |   |                             |
| ORGANISM              | Homo sapiens   | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; |                             |
| REFERENCE             | Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.   |   |                             |
| AUTHORS               | NIH-MGC http://mgc.nci.nih.gov/  |   |                             |
| TITLE                 | 1 (bases 1 to 928)   |   |                             |
| JOURNAL               | Unpublished (1999)   |   |                             |
| COMMENT               | National Institutes of Health, Mammalian Gene Collection (MGC)<br>Contact: Robert Strausberg, Ph.D.<br>Email: cgabs-femail.nih.gov<br>Tissue Procurement: ATCC<br>CDNA Library Preparation: Ling Hong/Rubin Laboratory<br>CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)<br>DNA Sequencing by: Incyte Genomics, Inc.<br>Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:<br>http://image.llnl.gov<br>Plate: L1CM183 row: m column: 18<br>High quality sequence start: 28<br>High quality sequence stop: 756.<br>Location/Qualifiers |   |                             |
| FEATURES              | source   | 1..928  |                             |
|                       | /organism="Homo sapiens"   |   |                             |
|                       | /db_xref="taxon:9606"  |   |                             |
|                       | /clone_image="5019065"   |   |                             |
|                       | /clone_lib="NIH_MGC_42"  |   |                             |
|                       | /tissue_type="epithelioid carcinoma cell line"   |   |                             |
|                       | /lab_host="DH10B (phage-resistant)"  |   |                             |
|                       | /note="Organ: pancreas; Vector: pORF7; Site_1: XhoI;<br>Site_2: EcoRI; CDNA made by oligo-dT priming.<br>Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies).<br>Note: this is a NIH_MGC library."   |   |                             |
| BASE COUNT            | 161 a 301 c 309 g 157 t  |   |                             |
| ORIGIN                |  |   |                             |
| Query Match           | 53.7%  | Score 404.2;  | Length 928;                 |
| Best Local Similarity | 76.2%;   | Pred. No. 5.8e-89;  |                             |
| Matches 551,          | Conservative 0;  | Mismatches 163;   | Indels 9; Gaps 4;           |
| Oy                    | 21   | CCATCTTCGCGCACACTTCTCTGCTGCTCTTAANGTGTGTGCCACACTGCGG              | 80                          |
|                       |  |   |                             |
| Db                    | 82   | CCACCTCTCGCCCTTCTCTGCTGCTCTCTAAGAAGTGTGCGTACCACTGTGCC             | 141                         |
| Oy                    | 81   | GACACCGCTGACGTGCTGTGACACACCCAGTCCGACAGGGGGTACCGCTGTGT             | 140                         |
|                       |  |   |                             |
| Db                    | 142  | GACACCACTGACCTGCCCCCTGCGCACCTCCCGATGCCGTGGAGTAGTCCCTGTGT          | 201                         |
| Oy                    | 141  | GGATGCGTGTGCTGTGAATGTGTGACAGAGAGCTGGGGAGATCCTGTGCACACT            | 200                         |
|                       |  |   |                             |
| Db                    | 202  | GGATGCGTGTGCTGTGCTGCGCGGATATGTGACACGGCGCTGGGGAGACCTGTGCACACT      | 261                         |
| Oy                    | 201  | GCATGTCTGCGACCCAGCCAGGGCTGTGTGTGACCTGTGGGACAGCCCTGTGGGCCA         | 260                         |
|                       |  |   |                             |
| Db                    | 262  | CCAAGTCTGGAGCCAGCCAGGGCTGTGTTCTGCAGCCCGGAGACCGGTTGGACG            | 321                         |
| Oy                    | 261  | TGGGGCTGTGTCTCTTGATAGAGTACGCTACTGTAGAGTGAATGGCCGACGTA             | 320                         |
|                       |  |   |                             |

|            |   |   |                             |
|------------|---|---|-----------------------------|
| Db         | 322   | GGGGGCCCTGTGCTCTTGTGGACAGAGACACAGACAGCTGTGAGTGTAAGGAGCGCCGCTCTA   | 381                         |
| Oy         | 321   | CTTGATGACAGACCTTTAAACCAATTGACAGAGGCTCTGTGCCCTGTGATGACGTTGG        | 380                         |
|            |   |   |                             |
| Db         | 382   | TGCGGAAGGGGAGACCTTCCAGCCCACTGCAGACATCTGCCTGCCCTCGAGAGAGCGCG       | 441                         |
| Oy         | 381   | CTTCACCTGCTGCTGCTGTGACAGATGTGGGGCTGCCACGTGGAGTGGCCACG             | 440                         |
|            |   |   |                             |
| Db         | 442   | CTTCACCTGCTGCTGCTGTGACAGATGTGGGGCTGCCACGTGGAGTGGAGTACCCCA         | 501                         |
| Oy         | 441   | CCCCAAGAAATACAGGTGACAGAGAAAGTCTGCCCGAGTGGTATGTGACAGAGAGT          | 500                         |
|            |   |   |                             |
| Db         | 502   | ACCAGAGAGGTGTGAGAGTCTGTGGCAAGTGTGCTGAGTGTGGTGGGCAAGAGAG           | 561                         |
| Oy         | 501   | GACACGGCGATCCAGCTCCACAGCGCGCAAGACACCACTTTTGTGCCCTTTGACT--         | 558                         |
|            |   |   |                             |
| Db         | 562   | GGGAGCTGTGGAGACCACTTACACAGCCCAAGAGACCAATTTGTGCTGTCTCTG            | 621                         |
| Oy         | 559   | CGTGCCCTGCTGATGCTGCTCTTGTCCAAATTTGAGACAGCGCTGGGG--CCCTGCTCAAC     | 617                         |
|            |   |   |                             |
| Db         | 622   | CCATGCCCACTGTGTCTCCCTGACCAAGATGAGACAGCGCTTTGGGGAACCTGCTCCAC       | 681                         |
| Oy         | 618   | CACCTGTGGG--CTGGGCAATAGCCACCGAGTGTCCCAACCAACCAAGTTCGCCAAC         | 675                         |
|            |   |   |                             |
| Db         | 682   | CACATGTGGGGCTGGGATTTGACACCGGGGTGTCCACAACAACCAAGATTTCTGCCACTG      | 741                         |
| Oy         | 676   | GAGATCCAA---GCGCGCTGTGTGTGCCACAGCTTCCTGTGACGACGACGACGACAG         | 731                         |
|            |   |   |                             |
| Db         | 742   | GAGACCCAAAGCGCGCTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG         | 801                         |
| Oy         | 732   | CTC 734   |                             |
|            |   |   |                             |
| Db         | 802   | GTC 804   |                             |
| RESULT 11  | Bt1825652   | 888 bp  | mRNA linear EST 04-OCT-2001 |
| LOCUS      | Bt1825652   | 60307263f1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5164614 5'   |                             |
| DEFINITION | Bt1825652   | mRNA sequence.  |                             |
| ACCESSION  | Bt1825652.1   | GI:15937202   |                             |
| VERSION    | Bt1825652.1   | GI:15937202   |                             |
| KEYWORDS   | EST.  |   |                             |
| SOURCE     | human.  |   |                             |
| ORGANISM   | Homo sapiens  | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; |                             |
| REFERENCE  | Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  |   |                             |
| TITLE      | 1 (bases 1 to 888)  |   |                             |
| AUTHORS    | NIH-MGC http://mgc.nci.nih.gov/   |   |                             |
| JOURNAL    | Unpublished (1999)  |   |                             |
| COMMENT    | National Institutes of Health, Mammalian Gene Collection (MGC)<br>Contact: Robert Strausberg, Ph.D.<br>Email: cgabs-femail.nih.gov<br>Tissue Procurement: Life Technologies, Inc.<br>CDNA Library Preparation: Life Technologies, Inc.<br>CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)<br>DNA Sequencing by: Incyte Genomics, Inc.<br>Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:<br>http://image.llnl.gov<br>Plate: LLML1408 row: n column: 07<br>High quality sequence stop: 87 |   |                             |



Db 422 TGCCCTGCTGTCCTCCCTGCCCCAGATGAGCAGCGGCTGGGACCTGCTCGACCACT 481  
 QY 623 GTGGGCTGGCATATGACCCAGAGTGTCTCCACAGACCAATTCGCAACTGAGATCC 682  
 Db 482 GTGGGCTGGGATGCGACCGGGGTGTCCACCAACAGCGTTCGCGAGCTGAGAACCC 541  
 QY 683 AACGCGGCTGTGTCTGCGCAGACCTGCTGGCAGCCAGCAGCAGCTCATGGAACA 742  
 Db 542 AGCGCGGCTGTGCTGTCTGCGCAGGCGCTGCGCAGGCTGCGAGTCCACAAACA 601  
 QY 743 GTGCTTCTA 752  
 Db 602 GTGCTTCTA 611

RESULT 13  
 BQ952960 1022 bp mRNA linear EST 21-AUG-2002  
 LOCUS AGENCOURT\_8931684 NIH\_MGC\_40 Homo sapiens cDNA clone IMAGE:6484123  
 DEFINITION 5', mRNA sequence.

ACCESSION BQ952960  
 VERSION BQ952960.1 GI:22368438  
 KEYWORDS EST.

SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 1022)  
 NIH-MGC http://mgs.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1995)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cga@pds-remail.nih.gov  
 Tissue Procurement: DCTD/DRP  
 cDNA Library Preparation: Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: L1CM2669 row: m column: 20  
 High quality sequence stop: 482.

FEATURES  
 source  
 Location/Qualifiers  
 1..1022

/organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone IMAGE:6484123  
 /clone\_11b="NIH\_MGC\_40"  
 /tissue\_type="carcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: prostate; Vector: pORF7; Site\_1: XhoI;  
 Site\_2: EcoRI; cDNA made by oligo-dT priming.  
 directionally cloned into EcoRI/XhoI sites using the  
 following 5' adaptor: GGCACAG(G). Library constructed by  
 Ling Hong in the laboratory of Gerald M. Rubin (University  
 of California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies).  
 Note: this is a NIH-MGC Library."

BASE COUNT 129 a 305 c 457 g 131 t  
 ORIGIN

Query Match 49.8%; Score 375.2; DB 14; Length 1022;  
 Best Local Similarity 78.1%; Pred. No. 8; le-82;  
 Matches 467; Conservative 0; Mismatches 123; Indels 8; Gaps 1;

QY 132 GGTGTGTAAGTGTGTGACGAGGCTGGGGAGTGTGACACCACTGCTGTCGCG 211  
 Db 1 GGTGTGCGGGGTATGTGACGCGGGGTGGGGAGCCGCGACCACTCCAGCTGTGCG 60  
 QY 212 ACCGACGACGAGGCTGTGTTGTGACGCGGGAGGCGCTGGCGCATGGGCTGTGT 271  
 Db 61 ACCGACGACGAGGCTGTGTTGTGACGCGGGAGGCGCGTGTGCGCGGGCGCTGT 120

QY 272 GTCTCT-----TGATGAGAGTACGCTGATGAGTGAATGCGCCGACGTACT 323  
 Db 121 GCTCTGTAAAGAGAGGAGGAGACAGAGAGCTGATGAGTGAAGGCGCCGCTGTATCG 180  
 QY 324 GGATGAGAGACCTTAAACCAATTGCAAGGCTCTGTGCGCGTGTGATGACGGGCTT 383  
 Db 181 GGAAGGAGAGACTTCCAGCGCCCACTGCAAGATCGCTGCGCGTGTGAGAGAGGCGCTT 240  
 QY 384 CACCTGCTGCGGCTGTGCTGAGTGAAGTGTGCGGCTGCGCAAGTGGAGTGGCCACGCCC 443  
 Db 241 CACCTGCGTGGCGCTGTGCTGAGAGATGTGGGCTGCGCAAGTGGAGTGGCCACGCCC 300  
 QY 444 CAGAGAAATPACAGTGTGCCAGGAAGAGTGTGCCCGAGTGGTATGACCAAGGAGTAC 503  
 Db 301 CAGAGAGGTGACAGGTCTGTGCAAGTGTGCTGAGTGTGCGCGCAAGGAGGCGG 360  
 QY 504 ACCGCGATGCCAGGCTCCAGGCGGCAAGACCACTTGTGCGCTGTGACACCTGCG 563  
 Db 361 ACTGGGAGCCAGCGCCCTTCCAGCCCAAGAGACCCAGTTTCTGTGCTCTTCCCT 420  
 QY 564 CTCTGTATGCTCTCTGTCCTCCAAATTTGAGAGACAGCTGGGCGCCCTGCAACACCTG 623  
 Db 421 GCCCCTGTGTGCTCTGCTGCGCAGATGAGACAGCGCTGGGACCTGCTGACACCTG 480  
 QY 624 TGGGCTGGCATATGACCCAGGAGTGTCCACCAAGACCAATTCGCAACTGAGATCCA 683  
 Db 481 TGGGCTGGCATATGACCCAGGAGTGTCCACCAAGACCAATTCGCAACTGAGATCCA 540  
 QY 684 ACCGCGCTGTGTGCTGCGCAGACCTGCTGGCAGCAGGAGGCCACACTGATGAGAC 741  
 Db 541 GCGCGCTGTGTGCTGCGCAGGCGCTGCGCAGGAGGCCACACTGATGAGAC 598

RESULT 14  
 LOCUS BG928868 620 bp mRNA linear EST 06-NOV-2001  
 DEFINITION HNC57-1-D9 R HNC (Human Normal Cartilage) Homo sapiens cDNA, mRNA  
 sequence.  
 ACCESSION BG928868  
 VERSION BG928868.1 GI:14323391  
 KEYWORDS EST.

SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 620)  
 Kumar S., Connor J.R., Dadds R.A., Halsey W., Van Horn M., Mao J.,  
 Sathe G., Mul P., Agarwal P., Badger A.M., Lee J.C., Gowen M. and  
 Lark M.W.

TITLE Identification and initial characterization of 5000 expressed  
 sequenced tags (ESTs) each from adult human normal and  
 osteoarthritic cartilage cDNA libraries  
 JOURNAL Osteoarthr. Cartil. 9 (7), 641-653 (2001)  
 MEDLINE 21482651  
 COMMENT Contact: Sanjay Kumar  
 UW2109

7105 Swedeland Road, P.O. Box 1539, King of Prussia, PA 19406, USA  
 Tel: 610-270-7245  
 Fax: 610-270-5598  
 Email: sanjay\_kumar-1@gsf.com  
 Seq primer: 17.

FEATURES  
 source  
 Location/Qualifiers  
 1..620

/organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_11b="HNC (Human Normal Cartilage)"  
 /tissue\_type="cartilage"  
 /lab\_host="E.coli DH10 B"  
 /note="Vector: pSPORT 1; Site\_1: SalI; Site\_2: NotI;  
 Directional"

BASE COUNT 97 a 218 c 207 g 98 t  
 ORIGIN







GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 15:19:16 ; Search time 4746.55 Seconds

(without alignments)  
10472.371 Million cell updates/sec

Title: US-10-010-408-1

Perfect score: 1708

Sequence: 1 GAGCGTCTGATCTCAGAG.....GCTTAGAATAACACCCCAA 1708

Scoring table: OLIGO\_NUC

Searched: 2054640 seqs, 14551402878 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl:\*

1: gb\_ba:\*

2: gb\_htg:\*

3: gb\_in:\*

4: gb\_om:\*

5: gb\_ov:\*

6: gb\_pat:\*

7: gb\_ph:\*

8: gb\_pl:\*

9: gb\_pr:\*

10: gb\_ro:\*

11: gb\_sts:\*

12: gb\_sy:\*

13: gb\_un:\*

14: gb\_vl:\*

15: em\_ba:\*

16: em\_fun:\*

17: em\_hum:\*

18: em\_in:\*

19: em\_mu:\*

20: em\_om:\*

21: em\_or:\*

22: em\_ov:\*

23: em\_pat:\*

24: em\_ph:\*

25: em\_pl:\*

26: em\_ro:\*

27: em\_sts:\*

28: em\_un:\*

29: em\_vl:\*

30: em\_htg\_hum:\*

31: em\_htg\_inv:\*

32: em\_htg\_other:\*

33: em\_htg\_mus:\*

34: em\_htg\_pln:\*

35: em\_htg\_rnd:\*

36: em\_htg\_mam:\*

37: em\_htg\_vrt:\*

38: em\_sy:\*

39: em\_htgo\_hum:\*

40: em\_htgo\_mus:\*

41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID       | Description        |
|------------|-------|-------------|--------|-------------|--------------------|
| 1          | 611   | 35.8        | 1741   | 10 AF259981 | AF259981 Rattus no |
| 2          | 382   | 22.4        | 137964 | 2 AC126895  | AC126895 Rattus no |
| 3          | 369   | 21.6        | 226303 | 2 AC095418  | AC095418 Rattus no |
| c          | 166   | 9.7         | 137964 | 2 AC126895  | AC126895 Rattus no |
| 4          | 90    | 5.3         | 1734   | 6 AR210324  | AR210324 Sequence  |
| c          | 90    | 5.3         | 1734   | 6 AR210325  | AR210325 Sequence  |
| 6          | 90    | 5.3         | 1734   | 10 AF100778 | AF100778 Mus muscu |
| 7          | 90    | 5.3         | 61072  | 10 AL731698 | AL731698 Mouse DNA |
| 8          | 90    | 5.3         | 216757 | 2 AL669906  | AL669906 Mus muscu |
| 9          | 80    | 4.7         | 1739   | 10 AF126063 | AF126063 Mus muscu |
| 10         | 80    | 4.7         | 738    | 6 AR210337  | AR210337 Sequence  |
| 11         | 32    | 1.9         | 841    | 6 AR210338  | AR210338 Sequence  |
| 12         | 32    | 1.9         | 1266   | 6 AX076919  | AX076919 Sequence  |
| 13         | 32    | 1.9         | 1266   | 6 AX464186  | AX464186 Sequence  |
| 14         | 32    | 1.9         | 1266   | 6 AF083500  | AF083500 Homo sapi |
| 15         | 32    | 1.9         | 1283   | 6 AR210322  | AR210322 Sequence  |
| 16         | 32    | 1.9         | 1293   | 6 AR210323  | AR210323 Sequence  |
| c          | 32    | 1.9         | 1293   | 6 AR210323  | AR210323 Sequence  |
| 17         | 32    | 1.9         | 1309   | 9 AF074604  | AF074604 Homo sapi |
| 18         | 32    | 1.9         | 1427   | 9 AF100780  | AF100780 Homo sapi |
| 19         | 32    | 1.9         | 1450   | 9 BC017782  | BC017782 Homo sapi |
| 20         | 32    | 1.9         | 1450   | 9 AL139352  | AL139352 Human DNA |
| 21         | 27    | 1.6         | 51     | 6 AR210371  | AR210371 Sequence  |
| 22         | 27    | 1.6         | 51     | 6 AX076923  | AX076923 Sequence  |
| 23         | 27    | 1.6         | 51     | 6 AC015962  | AC015962 Homo sapi |
| 24         | 23    | 1.3         | 145540 | 2 AC108665  | AC108665 Rattus no |
| 25         | 23    | 1.3         | 145897 | 2 AP002393  | AP002393 Homo sapi |
| c          | 23    | 1.3         | 176107 | 2 AC090246  | AC090246 Homo sapi |
| 26         | 23    | 1.3         | 176107 | 2 AC105227  | AC105227 Homo sapi |
| c          | 23    | 1.3         | 185568 | 9 AB057676  | AB057676 Oryzias l |
| 28         | 23    | 1.3         | 11621  | 5 AC010446  | AC010446 Homo sapi |
| 29         | 22    | 1.3         | 11621  | 5 AC091173  | AC091173 Homo sapi |
| c          | 30    | 1.3         | 142142 | 2 AC110904  | AC110904 Mus muscu |
| 31         | 22    | 1.3         | 149483 | 2 AC116351  | AC116351 Homo sapi |
| 32         | 22    | 1.3         | 180464 | 9 AC067881  | AC067881 Homo sapi |
| c          | 33    | 1.3         | 186676 | 2 AL646068  | AL646068 Ralstonia |
| 34         | 22    | 1.3         | 200050 | 1 AC119914  | AC119914 Mus muscu |
| c          | 35    | 1.2         | 66908  | 2 AC112856  | AC112856 Rattus no |
| 36         | 21    | 1.2         | 150944 | 2 AC122102  | AC122102 Rattus no |
| c          | 37    | 1.2         | 154881 | 2 AC103349  | AC103349 Mus muscu |
| 38         | 21    | 1.2         | 167691 | 2 AC114696  | AC114696 Rattus no |
| c          | 39    | 1.2         | 173702 | 2 AC107112  | AC107112 Rattus no |
| 40         | 21    | 1.2         | 179236 | 2 AL671882  | AL671882 Mouse DNA |
| 41         | 21    | 1.2         | 181343 | 10 AL831741 | AL831741 Mus muscu |
| c          | 42    | 1.2         | 204937 | 2 AC073717  | AC073717 Mus muscu |
| 43         | 21    | 1.2         | 215105 | 2 AL672068  | AL672068 Mouse DNA |
| 44         | 21    | 1.2         | 215105 | 2 AL672068  | AL672068 Mouse DNA |
| 45         | 21    | 1.2         | 217917 | 10 AL672068 | AL672068 Mouse DNA |

## ALIGNMENTS

RESULT 1  
AF259981  
LOCUS  
DEFINITION Rattus norvegicus CCN family protein COP-1 (COP-1) mRNA, complete  
CDs  
ACCESSION AF259981  
VERSION AF259981.1 GI:7739780  
KEYWORDS  
SOURCE Rattus norvegicus  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
REFERENCE 1 (bases 1 to 1741)  
Zhang, R., Averboukh, L., Zhu, W., Zhang, H., Jo, H., Dempsey, P. J.,

**TITLE** Coffey,R.J., Pardee,A.B. and Liang,P.  
 Identification of rCop-1, a new member of the CCN protein family,  
 as a negative regulator for cell transformation  
**JOURNAL** Mol. Cell. Biol. 18 (10), 6131-6141 (1998)  
**MEDLINE** 98414629  
**PUBMED** 9742130  
**REFERENCE** 2 (bases 1 to 1741)  
**AUTHORS** Liang,P.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (24-APR-2000) Cell Biology, Vanderbilt- Ingram Cancer  
 Center, 649 MRB II, Nashville, TN 37232, USA  
**FEATURES** Location/Qualifiers  
 source 1..1741  
 gene /organism="Rattus norvegicus"  
 CDS /db\_xref="taxon:10116"  
 1..1741  
 /gene="Cop-1"  
 262..1014  
 /note="secreted protein"  
 /protein\_id="AF69011.1"  
 /product="CCN family protein Cop-1"  
 /translation="MRGSPILRLATSPGLLSMVCADLCRPTCTPMTPPQCGVP  
 LVLDGCGCKVCAKRLTSCGHLHVCPEQGLVCGAGPGGAGVCLDEDDGCEV  
 NGRRLDGETFKPNCRLVLCRDGDEFTCLBLSEDTLPSPCCPRKRIQVPGKCP  
 WYCDQGVTPAIORSAOQHLSALVTPASADAMPWPMWSTAMPSCSTGIGIATRVSN  
 ONRFQLEIQRRLCPPLARSHSNMSAF"  
**BASE COUNT** 386 a 491 c 480 g 384 t  
**ORIGIN**  
**Query Match** 35.8%; Score 611; DB 10; Length 1741;  
**Best Local Similarity** 99.4%; Pred. No. 0;  
**Matches** 861; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 QY 665 GCAGCCAGCTGGAGATGCCACGCCCCCAAGAAATACAGAGTTCGCTCC 724  
 DB 678 GCAGCCAGCTGGAGATGCCACGCCCCCAAGAAATACAGAGTTCGCTCC 737  
 QY 725 CGAGTGGATATGACAGGAGATGACACGCGGATGACAGCGCTCCACGGCGCAAGACA 784  
 DB 738 CGAGTGGATATGACAGGAGATGACACGCGGATGACAGCGCTCCACGGCGCAAGACA 797  
 QY 785 CCAACTTTTGCCTTGTCTACCTCTGCTGCTGATGCTCTTGTCCAAATTGAGACAC 844  
 DB 798 CCAACTTTTGCCTTGTCTACCTCTGCTGCTGATGCTCTTGTCCAAATTGAGACAC 857  
 QY 845 AGCGTGGGCCCCCTGTCACACCCGCTGGGCGATACACCCGAGTTCACACA 904  
 DB 858 AGCGTGGGCCCCCTGTCACACCCGCTGGGCGATACACCCGAGTTCACACA 917  
 QY 905 GAACCATTCCTGCACTGGAGATGACAGCGCCCTGTGTCTGCGACAGACCTGCTGGC 964  
 DB 918 GAACCATTCCTGCACTGGAGATGACAGCGCCCTGTGTCTGCGACAGACCTGCTGGC 977  
 QY 965 AGCCAGAGCCACAGCTCATGAAAGTCTTTTAAAGCCAACTGGGATGCGATACA 1024  
 DB 978 AGCCAGAGCCACAGCTCATGAAAGTCTTTTAAAGCCAACTGGGATGCGATACA 1037  
 QY 1025 GGGCCCTGCATCTCTACAGCAATACCTAGACAGCCCTGAGCTGCTGATGCTC 1084  
 DB 1038 GGGCCCTGCATCTCTACAGCAATACCTGAGACAGCCCTGAGCTGCTGATGCTC 1097  
 QY 1085 TTCTCATGCTCTTGAGCTGAGTAACTGCTGCTGATGCTGCTGATGAGGACATG 1144  
 DB 1098 TTCTCATGCTCTTGAGCTGAGTAACTGCTGCTGATGCTGCTGATGAGGACATG 1157  
 QY 1145 AGCGATCCCTGCTCTCTGAGGTAGCGAGAGCTGACACAGCTTCAGTCTCTGTTT 1204  
 DB 1158 AGCGATCCCTGCTCTCTGAGGTAGCGAGAGCTGACACAGCTTCAGTCTCTGTTT 1217  
 QY 1205 AGCGTGAATTCCTGCTCTGCTGCTGCTCAAAAACATCCCTGTCACAAAAGGACA 1264

DB 1218 AGCGTGAATTCCTGCTCTGCTGCTGCTCAAAAACATCCCTGTCACAAAAGGACA 1277  
 QY 1265 ACCAAAAGACCTTTAACTAGGCTATPACTGGGCAAACTGGCCACCGTGGCGGATTA 1324  
 DB 1278 ACCAAAAGACCTTTAACTAGGCTATPACTGGGCAAACTGGCCACCGTGGCGGATTA 1337  
 QY 1325 AGGTCAATGTATGAGCAGACAGACAGATGCTGTAACCTTCAGTTCCTTGTGACTT 1384  
 DB 1338 AGGTCAATGTATGAGCAGACAGACAGATGCTGTAACCTTCAGTTCCTTGTGACTT 1397  
 QY 1385 CTGTATGCTTTGTCCTCAAAAGATGATGATGACTGTAAGTACTTCTCCCTGACTGAG 1444  
 DB 1398 CTGTATGCTTTGTCCTCAAAAGATGATGATGACTGTAAGTACTTCTCCCTGACTGAG 1457  
 QY 1445 AACACCCCTGCTGCTGGGAGATGATGAGGAGCAATCTCTGTGACATGAAAGAGT 1504  
 DB 1458 AACACCCCTGCTGCTGGGAGATGATGAGGAGCAATCTCTGTGACATGAAAGAGT 1517  
 QY 1505 AATCACACTGCTCTTAAAGAAATTCCT 1530  
 DB 1518 AATCACACTGCTCTTAAAGAAATTCCT 1543  
**RESULT 2**  
**AC126895** 137964 bp DNA linear HTG 24-JUL-2002  
**LOCUS** Rattus norvegicus clone CH230-301B, \*\*\* SEQUENCING IN PROGRESS  
**DEFINITION** \*\*\* 49 unordered pieces.  
**ACCESSION** AC126895  
**VERSION** AC126895.1 GI:21724040  
**KEYWORDS** HTG; HTGS; PHASE1.  
**SOURCE** Rattus norvegicus.  
**ORGANISM** Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
**REFERENCE** 1 (bases 1 to 137964)  
 Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,  
 Alsbrooks,S.L., Amaralunga,H.C., Are,J.R., Ayele,M., Banks,T.,  
 Barbato,J., Benton,J., Blomberg,K., Blankenburg,K., Bonin,D.,  
 Bouck,J., Bowie,S., Briley,M., Brown,E., Brown,M., Bryant,N.P.,  
 Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,  
 Cartron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,  
 Chen,G., Chen,R., Chen,Z., Chowdhury,I., Christopoulos,C.,  
 Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,  
 Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,  
 Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,  
 Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,  
 Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,  
 Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,  
 Gabisi,J., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,  
 Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,  
 Harris,C., Harris,K., Hart,M., Haylak,P., Hawes,A., Hernandez,J.,  
 Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,  
 Homsi,E., Howard,S., Huber,J., Huliy,S., Hune,U., Jackson,L.E.,  
 Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,  
 Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,  
 Kravovic,J., Kureishi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,  
 Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louised,H.,  
 Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,  
 Maheshwari,M., Mapa,P., Martin,R., Martindale,A., Martinez,E.,  
 Massey,E., Mayhew,E., McLeod,M.P., Meador,M., Mel,G., Metzger,M.,  
 Miner,G., Miner,Z., Mitchell,T., Mohabat,K., Morgan,M., Morris,S.,  
 Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,  
 Nguyen,N., Nickerson,E., Nwokoko,S., Oguh,M., Okunolu,G.,  
 Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,  
 Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,  
 Rives,M., Rojas,A., Rojudo-Kan,I., Rolfe,M., Ruiz,S., Saverly,G.,  
 Scherer,S., Scott,G., Shen,H., Shooshari,M., Sisson,L.,  
 Sodergren,E., Sonalike,T., Sparks,A., Stanley,H., Stone,H.,  
 Sutton,A., Svatek,A., Taber,P., Tamerisa,A., Tamerisa,K., Tang,H.,  
 Tansey,J., Taylor,C., Taylor,T., Tellier,B., Thomas,N., Thomas,S.,

Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,O.,  
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,  
Williams,G., Williamson,A., Mieczyslaw,R., Woodson,S., Worley,K.,  
Wu,C., Wu,Y., Wu,Y.F., Zhou,Y., Zorrilla,S., Nelson,D.,  
Weinstock,G. and Gibbs,R.  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
Center: Baylor College of Medicine  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
Project Information  
Center project name: GZHG  
Center clone name: CH230-301E4  
Summary Statistics  
Sequencing vector: Plasmid;  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 89474 bases at least Q40  
Consensus quality: 93422 bases at least Q30  
Consensus quality: 96506 bases at least Q20  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 49 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
1 1056: contig of 1056 bp in length  
\* 1057 1156: gap of unknown length  
\* 1157 2335: contig of 1179 bp in length  
\* 2336 2435: gap of unknown length  
\* 2436 3442: contig of 1007 bp in length  
\* 3443 3542: gap of unknown length  
\* 3543 5081: contig of 1539 bp in length  
\* 5082 5181: gap of unknown length  
\* 5182 6424: contig of 1243 bp in length  
\* 6425 6524: gap of unknown length  
\* 6525 7814: contig of 1290 bp in length  
\* 7815 7913: gap of unknown length  
\* 7913 9153: contig of 1239 bp in length  
\* 9154 9253: gap of unknown length  
\* 9254 10445: contig of 1192 bp in length  
\* 10446 10545: gap of unknown length  
\* 10546 11771: contig of 1226 bp in length  
\* 11772 11871: gap of unknown length  
\* 11872 13240: contig of 1369 bp in length  
\* 13241 13340: gap of unknown length  
\* 13341 14359: contig of 1019 bp in length  
\* 14360 14459: gap of unknown length  
\* 14460 15651: contig of 1192 bp in length  
\* 15652 15751: gap of unknown length  
\* 15752 17494: contig of 1743 bp in length  
\* 17495 17594: gap of unknown length  
\* 17595 18679: contig of 1085 bp in length  
\* 18680 18779: gap of unknown length

18780 20681: contig of 1902 bp in length  
\* 20682 20782: gap of unknown length  
\* 20783 22118: contig of 1337 bp in length  
\* 22119 22218: gap of unknown length  
\* 22219 23578: contig of 1360 bp in length  
\* 23579 23678: gap of unknown length  
\* 23679 25423: contig of 1745 bp in length  
\* 25424 25523: gap of unknown length  
\* 25524 27808: contig of 2285 bp in length  
\* 27809 27908: gap of unknown length  
\* 27909 30272: contig of 2364 bp in length  
\* 30273 30372: gap of unknown length  
\* 30373 33091: contig of 2719 bp in length  
\* 33092 33191: gap of unknown length  
\* 33192 35777: contig of 2586 bp in length  
\* 35778 35877: gap of unknown length  
\* 35878 36908: contig of 1031 bp in length  
\* 36909 37008: gap of unknown length  
\* 37009 38930: contig of 1922 bp in length  
\* 38931 39030: gap of unknown length  
\* 39031 41210: contig of 2180 bp in length  
\* 41211 41310: gap of unknown length  
\* 41311 42937: contig of 1627 bp in length  
\* 42938 43037: gap of unknown length  
\* 43038 46120: contig of 3083 bp in length  
\* 46121 46220: gap of unknown length  
\* 46221 48217: contig of 1997 bp in length  
\* 48218 48317: gap of unknown length  
\* 48319 50433: contig of 2116 bp in length  
\* 50434 50533: gap of unknown length  
\* 50534 52880: contig of 2347 bp in length  
\* 52881 52980: gap of unknown length  
\* 52981 55194: contig of 2214 bp in length  
\* 55195 55294: gap of unknown length  
\* 55295 57482: contig of 2188 bp in length  
\* 57483 57582: gap of unknown length  
\* 57583 60781: contig of 3199 bp in length  
\* 60782 60881: gap of unknown length  
\* 60882 62699: contig of 1718 bp in length  
\* 62699 62700: gap of unknown length  
\* 62700 65203: contig of 2504 bp in length  
\* 65204 66914: gap of unknown length  
\* 66914 67014: contig of 1611 bp in length  
\* 67015 70839: gap of unknown length  
\* 70840 70939: contig of 3825 bp in length  
\* 70940 75139: gap of unknown length  
\* 75140 75239: contig of 4200 bp in length  
\* 75240 78296: gap of unknown length  
\* 78297 78396: contig of 3057 bp in length  
\* 78397 81338: gap of unknown length  
\* 81339 83238: contig of 4742 bp in length  
\* 83239 88204: gap of unknown length  
\* 88205 88304: contig of 4966 bp in length  
\* 88305 92238: gap of unknown length  
\* 92239 92338: contig of 3934 bp in length  
\* 92339 97439: gap of unknown length  
\* 97440 97439: contig of 5001 bp in length  
\* 97440 103534: gap of unknown length  
\* 103535 103634: contig of 6095 bp in length  
\* 103635 107080: gap of unknown length  
\* 107081 107180: contig of 3446 bp in length  
\* 107181 115201: gap of unknown length  
\* 115202 115301: contig of 8021 bp in length  
\* 115302 121237: gap of unknown length  
\* 121238 121337: contig of 5936 bp in length  
\* 121338 127910: gap of unknown length  
\* 127911 128010: contig of 6573 bp in length  
\* 128011 137964: gap of unknown length  
\* 137964 13964: contig of 9954 bp in length.  
FEATURES  
source  
Location/Qualifiers  
1. 137964  
/organism="Rattus norvegicus"  
/db\_xref="taxon:10116"





Alshrook, S. I., Amaralunga, H. C., Are, J. R., Ayele, M., Banks, T., Barabara, J., Benton, J., Bimaga, K., Blankenburg, K., Bonini, D., Bouck, J., Boyle, S., Brivela, M., Brown, E., Brown, M., Bryant, N. P., Bunyah, C., Butch, P., Burkett, C., Burrell, K. L., Byrd, N. C., Carron, T. F., Carter, M., Cavazos, S. R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhury, I., Christopoulos, I., Cleveland, C. D., Cox, C., Coyle, M. D., Dathorne, S. R., David, R., Davila, M. L., Davis, C., Davy-Carroll, L., Decker, D. A., Delaney, K. R., Delgado, O., Dem, A. L., Ding, Y., Dinh, H. H., Douthwaite, K. J., Draper, H., Dugan-Rocha, S., Durbin, K. J., Earhart, C., Ferraguto, D., Edwards, C. C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frintz, P., Gabis, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J. H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogue, M., Holmway, C., Hollins, B., Homsl, F., Howard, S., Huber, J., Huliy, S., Hume, J., Jackson, L. E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Kallison, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovac, C., Kretovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L. C., Lewis, L. J., Li, Z., Lichtarge, O., Iken, C., Liu, J., Liu, M., Louised, H., Locado, R. J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapue, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mahoney, E., McLeod, M. P., Meador, M., Mei, G., Melker, M., Miner, G., Miner, Z., Mitchell, T., Monabbe, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenwo, S., Oguh, M., Okunolu, G., Orquyue, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L., Quiles, M., Ren, Y., Rivas, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savary, G., Scherer, S., Scott, G., Shen, H., Shooshitay, N., Sisson, I., Soederren, E., Sonalke, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, K., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalob, D., Vanson, R., Wang, O., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczek, R., Wood, S., Worley, K., Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorilla, S., Nelson, D., Weinstein, G., and Gibbs, R.

Direct Submission  
Unpublished  
2 (bases 1 to 137964)  
Worley, K. C.

Direct Submission  
Submitted (10-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 137964)  
Worley, K. C.

Direct Submission  
Submitted (24-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
Project Information  
Center project name: G2HG  
Center clone name: CH230-301E4  
Sequencing vector: Plasmid  
Summary Statistics  
Chemistry: Dye-terminator Big Dye 3.00% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 89474 bases at least Q40  
Consensus quality: 93422 bases at least Q30  
Consensus quality: 96506 bases at least Q20

NOTE: Estimated insert size may differ from sequence length  
(see [http://www.hgsc.bcm.tmc.edu/dogs/genbank\\_draft\\_c.data.html](http://www.hgsc.bcm.tmc.edu/dogs/genbank_draft_c.data.html)).  
NOTE: This is a working draft sequence. It currently  
consists of 49 contigs. The true order of the pieces

is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 1056: contig of 1056 bp in length  
1057 1156: gap of unknown length  
1157 2235: contig of 1179 bp in length  
2235 2336: gap of unknown length  
2336 2436: contig of 1007 bp in length  
2436 3443: gap of unknown length  
3443 3583: contig of 1539 bp in length  
3583 5081: gap of unknown length  
5081 5182: contig of 1243 bp in length  
5182 6425: gap of unknown length  
6425 7814: contig of 1250 bp in length  
7814 7915: gap of unknown length  
7915 9253: contig of 1239 bp in length  
9253 9254: gap of unknown length  
9254 10445: contig of 1192 bp in length  
10445 10545: gap of unknown length  
10545 10546: gap of unknown length  
10546 11771: contig of 1226 bp in length  
11771 11871: gap of unknown length  
11871 13340: contig of 1369 bp in length  
13340 14359: gap of unknown length  
14359 14459: contig of 1019 bp in length  
14459 15651: gap of unknown length  
15651 15751: contig of 1192 bp in length  
15751 17494: gap of unknown length  
17494 17595: contig of 1743 bp in length  
17595 18779: gap of unknown length  
18779 20681: contig of 1085 bp in length  
20681 22118: gap of unknown length  
22118 22218: contig of 1337 bp in length  
22218 23578: gap of unknown length  
23578 23678: contig of 1360 bp in length  
23678 25423: gap of unknown length  
25423 25523: contig of 1745 bp in length  
25523 27808: gap of unknown length  
27808 27909: contig of 2285 bp in length  
27909 30373: gap of unknown length  
30373 30373: contig of 2364 bp in length  
30373 33091: gap of unknown length  
33091 33191: contig of 2719 bp in length  
33191 35777: gap of unknown length  
35777 35877: contig of 2586 bp in length  
35877 36908: gap of unknown length  
36908 37008: contig of 1031 bp in length  
37008 38930: gap of unknown length  
38930 39030: contig of 1922 bp in length  
39030 41210: gap of unknown length  
41210 41310: contig of 2180 bp in length  
41310 42937: gap of unknown length  
42937 43037: contig of 1627 bp in length  
43037 46120: gap of unknown length  
46120 48317: contig of 3083 bp in length  
48317 48317: gap of unknown length  
48317 50433: gap of unknown length  
50433 50533: contig of 2116 bp in length  
50533 52880: gap of unknown length  
52880 52980: contig of 2347 bp in length  
52980 55195: gap of unknown length  
55195 55294: contig of 2214 bp in length  
55294 57482: gap of unknown length  
57482 57582: contig of 2188 bp in length  
57582 60781: gap of unknown length  
60781 60882: contig of 3199 bp in length  
60882 62599: gap of unknown length  
62599 60882: contig of 1718 bp in length





aberrantly expressed in human colon tumors  
 Proc. Natl. Acad. Sci. U.S.A. 95 (25), 14717-14722 (1998)

JOURNAL  
 MEDLINE  
 99061933  
 PUBMED  
 9843955  
 REFERENCE  
 2 (bases 1 to 1734)  
 AUTHORS  
 Pennica, D.  
 TITLE  
 Direct Submission  
 Submitted (23-Oct-1998) Molecular Oncology, Genentech, Inc., 1 DNA  
 May, South San Francisco, CA 94080, USA

FEATURES  
 source  
 1. 1734  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /cell\_line="C57MG"  
 /cell\_type="epithelial"  
 /tissue\_type="mammary"  
 /note="transformed by Mnt-1"  
 1. 1734  
 /gene="Wisp2"  
 257. 1012  
 /gene="Wisp2"  
 /codon\_start=1  
 /product="connective tissue growth factor related protein  
 Wisp-2"  
 /protein\_id="AAC96320.1"  
 /db\_xref="GI:4028579"  
 /translation="MRGNPLIHLAISFLCITSMVYSLCPAPCAPMTPPQCPGV  
 LVLDGCGCVRCAKRLGSCDHLHYCDPSGAGCPGAPGRCVAVCLFEEDSCSEV  
 NGRRLDGEFFKPCRYLCRCDDGFFCLPLCSDDVRLPSMDCPRRIOVPGCCPE  
 WYCDQAWQAPRIOPSSAOGHLSALYTPASADGPCPNSTWAGFCSTTCGIGATRVS  
 NONRCEIQRICLSRPLASSHSGSNSTF"

BASE COUNT  
 355 a 491 c 495 g 393 t

ORIGIN

Query Match  
 Best Local Similarity 100.0%; Score 90; DB 10; Length 1734;  
 Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 410 AGTGTGTGACGAGAGCTGGGAGATCCTCGACACCTGCATGTCTCGACCCAGCCA 469  
 |||||||  
 DB 418 AGTGTGTGACGAGAGCTGGGAGATCCTCGACACCTGCATGTCTCGACCCAGCCA 477  
 |||||||

QY 470 GGGCTGTTTGTTCAGCTGGGCGACGCC 499  
 |||||||

DB 478 GGGCTGTTTGTTCAGCTGGGCGACGCC 507  
 |||||||

RESULT 8  
 AL731698 61072 bp DNA linear ROD 24-MAY-2002  
 LOCUS Mouse DNA sequence from clone RP23-161B3 on chromosome 2, complete  
 sequence.  
 ACCESSION AL731698  
 VERSION AL731698.10 GI:21214309  
 KEYWORDS HTG;  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 61072)  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT  
 Direct Submission  
 Submitted (24-MAY-2002) Wellcome Trust Sanger Institute, Hinxton,  
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
 humquerry@sanger.ac.uk  
 On May 25, 2002 this sequence version replaced gi:21213601.  
 During sequence assembly data is compared from overlapping clones.  
 Where differences are found these are annotated as variations  
 together with a note of the overlapping clone name. Note that the  
 variation annotation may not be found in the sequence submission  
 corresponding to the overlapping clone, as we submit sequences with  
 only a small overlap as described above.  
 This sequence was finished as follows unless otherwise noted: all

regions were either double-stranded or sequenced with an alternate  
 chemistry or covered by high quality data (i.e., phred quality >=  
 30); an attempt was made to resolve all sequencing problems, such  
 as compressions and repeats; all regions were covered by at least  
 one plasmid subclone or more than one M13 subclone; and the  
 assembly was confirmed by restriction digest. The following  
 abbreviations are used to associate primary accession numbers given  
 in the feature table with their source databases: Em., EMBL; SW.,  
 SWISSPROT; Tr., TREMBL; Mp., WormPEP; Information on the WormPEP  
 database can be found at  
<http://www.sanger.ac.uk/projects/C-elegans/wormpep> RP23-161B3 is  
 from the RP23 mouse PAC library  
 constructed by the group of Pieter de Jong.  
 For further details see <http://www.chori.org/bacpac/home.htm>  
 VECTOR: pBACe3.6.

FEATURES  
 source  
 Location/Qualifiers  
 1. 61072  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /chromosome="2"  
 /clone="RP23-161B3"  
 /clone\_lib="RP23-161B3"

BASE COUNT  
 15405 a 15368 c 15033 g 15266 t

ORIGIN

Query Match  
 Best Local Similarity 100.0%; Score 90; DB 10; Length 61072;  
 Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 410 AGTGTGTGACGAGAGCTGGGAGATCCTCGACACCTGCATGTCTCGACCCAGCCA 469  
 |||||||  
 DB 10029 AGTGTGTGACGAGAGCTGGGAGATCCTCGACACCTGCATGTCTCGACCCAGCCA 10088  
 |||||||

QY 470 GGGCTGTTTGTTCAGCTGGGCGACGCC 499  
 |||||||

DB 10089 GGGCTGTTTGTTCAGCTGGGCGACGCC 10118  
 |||||||

RESULT 9  
 AL669906 216757 bp DNA linear HTG 24-JUL-2002  
 LOCUS Mus musculus chromosome 2 clone RP23-217C2, \*\*\* SEQUENCING IN  
 PROGRESS \*\*\*, 21 unordered pieces.  
 ACCESSION AL669906  
 VERSION AL669906.5 GI:21955520  
 KEYWORDS HTG; HTGS\_PHASE1; HTGS\_CANCELLED.  
 SOURCE Mus musculus.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 216757)  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT  
 Direct Submission  
 Submitted (23-JUL-2002) Wellcome Trust Sanger Institute, Hinxton,  
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
 humquerry@sanger.ac.uk  
 On Jul 25, 2002 this sequence version replaced gi:1818193.  
 ----- Genome Center  
 Center: Wellcome Trust Sanger Institute  
 Center code: SC  
 Web site: <http://www.sanger.ac.uk>  
 Contact: humquerry@sanger.ac.uk  
 ----- Project Information  
 Center project name: bm217C2  
 ----- Summary Statistics  
 Assembly program: XAPP4; version 4.5  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Consensus quality: 21165 bases at least Q40  
 Consensus quality: 213403 bases at least Q30  
 Consensus quality: 214139 bases at least Q20  
 Insert size: 214757; sum-of-contigs  
 Insert size: 234243; 1.8% error; agarose-fp  
 Quality coverage: 6.29x in Q20 bases; sum-of-contigs Quality

Coverage: 5.90x in Q20 bases; agarose-1p

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 21 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 4000: contig of 4000 bp in length  
\* 4001 4100: gap of 100 bp  
\* 4101 15267: contig of 11167 bp in length  
\* 15268 15367: gap of 100 bp  
\* 15368 34970: contig of 19603 bp in length  
\* 34971 35070: gap of 100 bp  
\* 35071 39671: contig of 4601 bp in length  
\* 39672 39771: gap of 100 bp  
\* 39772 46360: contig of 6589 bp in length  
\* 46361 46460: gap of 100 bp  
\* 46461 49809: contig of 3349 bp in length  
\* 49810 49909: gap of 100 bp  
\* 49910 54509: contig of 4600 bp in length  
\* 54510 54609: gap of 100 bp  
\* 54610 65989: contig of 11380 bp in length  
\* 65990 66089: gap of 100 bp  
\* 66090 84635: contig of 18546 bp in length  
\* 84636 84735: gap of 100 bp  
\* 84736 87399: contig of 2664 bp in length  
\* 87400 87499: gap of 100 bp  
\* 87500 98601: contig of 11102 bp in length  
\* 98602 98701: gap of 100 bp  
\* 98702 105522: contig of 6821 bp in length  
\* 105523 105622: gap of 100 bp  
\* 105623 119773: contig of 14151 bp in length  
\* 119774 119873: gap of 100 bp  
\* 119874 134552: contig of 14679 bp in length  
\* 134553 134652: gap of 100 bp  
\* 134653 141202: contig of 6550 bp in length  
\* 141203 141302: gap of 100 bp  
\* 141303 145808: contig of 4506 bp in length  
\* 145809 145908: gap of 100 bp  
\* 145909 151572: contig of 5664 bp in length  
\* 151573 151672: gap of 100 bp  
\* 151673 169686: contig of 18014 bp in length  
\* 169687 169786: gap of 100 bp  
\* 169787 200095: contig of 30309 bp in length  
\* 200096 200195: gap of 100 bp  
\* 200196 213877: contig of 13682 bp in length  
\* 213878 213977: gap of 100 bp  
\* 213978 216757: contig of 2780 bp in length.

# FEATURES

## SOURCE

1. 216757  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/chromosome="2"  
/clone\_lib="RP23-217C2"  
/clone\_lib="RPCT-23"  
1. 4000  
/note="assembly\_fragment:03429  
clone\_end:sp6  
vector\_side:left"  
4101..15267  
/note="assembly\_fragment:00814  
fragment\_chain:1"  
15368..34970  
/note="assembly\_fragment:00812  
fragment\_chain:1"  
35071..39671  
/note="assembly\_fragment:03544  
fragment\_chain:1"  
39772..46360  
/note="assembly\_fragment:01438

misc\_feature fragment\_chain:1"  
46461..49809  
/note="assembly\_fragment:03410  
fragment\_chain:2"  
49910..54509  
/note="assembly\_fragment:00811  
fragment\_chain:2"  
54610..65989  
/note="assembly\_fragment:02217  
fragment\_chain:2"  
66090..84635  
/note="assembly\_fragment:03529  
fragment\_chain:2"  
84736..87399  
/note="assembly\_fragment:01019  
fragment\_chain:3"  
87500..98601  
/note="assembly\_fragment:01170  
fragment\_chain:3"  
98702..105522  
/note="assembly\_fragment:02829  
fragment\_chain:3"  
105623..119773  
/note="assembly\_fragment:01476  
fragment\_chain:4"  
119874..134552  
/note="assembly\_fragment:00505  
fragment\_chain:4"  
134653..141202  
/note="assembly\_fragment:01951  
fragment\_chain:4"  
141303..145808  
/note="assembly\_fragment:02637  
fragment\_chain:5"  
145909..151572  
/note="assembly\_fragment:02352  
fragment\_chain:5"  
151673..169686  
/note="assembly\_fragment:02207  
fragment\_chain:5"  
169787..200095  
/note="assembly\_fragment:02849  
fragment\_chain:6"  
200196..213877  
/note="assembly\_fragment:01052  
fragment\_chain:6"  
213978..216757  
/note="assembly\_fragment:01318  
clone\_end:r7  
vector\_side:right"  
BASE COUNT 56425 a 51573 c 51904 g 54852 t 2003 others  
ORIGIN  
Query Match 5.3%; Score 90; DB 2; Length 216757;  
Best Local Similarity 100.0%; Pred. No. 5.9e-38;  
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 410 AGGTGTGACGAGAGCTGGGGAGTCTGCGACACCTGCATGTCTGCGACCCAGCCA 469  
|||||  
DB 166450 AGTGTGTGACGAGAGCTGGGGAGTCTGCGACACCTGCATGTCTGCGACCCAGCCA 166509  
|||||  
OY 470 GGGCCTGGTTGTGACGCTGGGGAGGCCC 499  
|||||  
DB 166510 GGGCCTGGTTGTGACGCTGGGGAGGCCC 166539  
|||||  
RESULT 10  
AF126063 1739 bp mRNA linear ROD 12-OCT-1999  
LOCUS AF126063  
DEFINITION Mus musculus connective tissue growth factor-like protein precursor  
(Ctgfl) mRNA, complete cds.  
ACCESSION AF126063  
VERSION AF126063.1 GI:4337059

KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source

Mus musculus.  
Mus musculus  
Mammalia: Eutheria: Chordata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Rodentia: Sciurognathia: Muridae: Murinae: Mus.  
1 (bases 1 to 1739)  
Kumar, S., Hand, A.T., Connor, J.R., Dodds, R.A., Ryan, P.J., Trill, J.J., Fisher, S.M., Nuttall, M.E., Lipschutz, D.B., Zou, C., Hwang, S.M., Votta, B.J., James, I.E., Rietman, D.J., Gowen, M. and Lee, J.C.  
Identification and cloning of a connective tissue growth factor-like cDNA from human osteoblasts encoding a novel regulator of osteoblast functions  
J. Biol. Chem. 274 (24), 17123-17131 (1999)  
99287915  
10358067  
2 (bases 1 to 1739)  
Kumar, S. and Zou, C.  
Direct Submission  
Submitted (04-FEB-1999) Bone & Cartilage Biology, UW 2109, SmithKline Beecham, 709 Swedeland Rd., King of Prussia, PA 19406, USA

1. 1739  
Location/Qualifiers  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/tissue\_type="lung"  
1. 1739  
/gene="Ctgfl"  
242..997  
/note="Ctgfl"  
/note="similar to the Mus musculus WISP-2 protein encoded by the sequence presented in Genbank Accession Number AF100778; putative growth factor: CTGF-L; contains IGF binding (IGFBP), Von Willebrand Factor type C (VWC) repeat and thrombospondin type I (TSP1) domains; member of the CCN (CTGF/Cyrf61/Nov) family; lacks the fourth carboxy-terminal (CT) domain present in other members of the CCN family"  
/codon\_start=1  
/product="connective tissue growth factor-like protein precursor"  
/protein\_id="AA18058.1"  
/db\_xref="GI:4337060"  
/translation="MRGNPLHLAISFLCISMVYSQLEPCAPACPMTPPOCPGCV LVLDCGCCRCVARRLGEESCDLHVCDPSQGLVCGAGPSRGACLPFEEDGSCVE NSRRYLFETFPKNCVILCRCDGSGFTICLPSSEVDRLSPMCPRRRIIOVGRCQPE WYCDVAVMGPATDPSASAOCHOLSAIVTPASLPCPCPNMSTANGPCTTGGLIATRVSNONRQQLTORLCLSRPCLASRSHGSMNSAF"

BASE COUNT 375 a 480 c 489 g 395 t  
ORIGIN

Query Match 4.7%; Score 80; DB 10; Length 1739;  
Best Local Similarity 100.0%; Pred. No. 3.7e-32;  
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 420 CGAGAGCTGGGGAGTCTGCGACACATGTCGTGAGACCCAGCCAGGGCTGTGTT 479  
|||||  
Db 413 CGAGAGCTGGGGAGTCTGCGACACATGTCGTGAGACCCAGCCAGGGCTGTGTT 472  
|||||

OY 480 TGTACACCTGGGGAGCC 499  
|||||  
Db 473 TGTACACCTGGGGAGCC 492  
|||||

RESULT 11  
AR210337 738 bp DNA linear PAT 20-JUN-2002  
LOCUS  
DEFINITION Sequence 38 from patent US 6387657.  
ACCESSION AR210337  
VERSION AR210337.1 GI:21512542  
KEYWORDS  
SOURCE Unknown.

ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source

Unknown.  
Unclassified.  
1 (bases 1 to 738)  
Botstein, D.A., Cohen, R.L., Goddard, A.D., Gurney, A.L., Hillan, K.J., Lawrence, D.A., Levine, A.J., Penhale, D., Roy, M.Ann. and Wood, W.I.  
WISP polypeptides and nucleic acids encoding same  
Patent: US 6387657-A 38 14-MAY-2002;  
Location/Qualifiers  
1. 738  
/organism="unknown"

BASE COUNT 104 a 272 c 238 g 124 t  
ORIGIN

Query Match 1.9%; Score 32; DB 6; Length 738;  
Best Local Similarity 100.0%; Pred. No. 1.3e-05;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 375 GTACCCCTGCTGCTGATGGCTGTGCTGCTG 406  
|||||  
Db 115 GTACCCCTGCTGCTGATGGCTGTGCTGCTG 146  
|||||

RESULT 12  
AR210338 841 bp DNA linear PAT 20-JUN-2002  
LOCUS  
DEFINITION Sequence 39 from patent US 6387657.  
ACCESSION AR210338  
VERSION AR210338.1 GI:21512543  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 841)  
AUTHORS Botstein, D.A., Cohen, R.L., Goddard, A.D., Gurney, A.L., Hillan, K.J., Lawrence, D.A., Levine, A.J., Penhale, D., Roy, M.Ann. and Wood, W.I.  
WISP polypeptides and nucleic acids encoding same  
Patent: US 6387657-A 39 14-MAY-2002;  
Location/Qualifiers  
1. 841  
/organism="unknown"

BASE COUNT 124 a 297 c 280 g 140 t  
ORIGIN

Query Match 1.9%; Score 32; DB 6; Length 841;  
Best Local Similarity 100.0%; Pred. No. 1.3e-05;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 654 GAGATGTCGGCTGCCAGCTGGGACTGCC 685  
|||||  
Db 417 GAGATGTCGGCTGCCAGCTGGGACTGCC 448  
|||||

RESULT 13  
AX076919 1266 bp DNA linear PAT 22-FEB-2001  
LOCUS  
DEFINITION Sequence 31 from Patent WO0105836.  
ACCESSION AX076919  
VERSION AX076919.1 GI:13121575  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Primates: Catarrhini: Homiidae: Homo.  
1 (bases 1 to 1266)  
Botstein, D., Goddard, A., Gurney, A.L., Hillan, K.J., Roy, M.A. and Wood, W.I.  
Polypeptide compositions and methods for the treatment of tumors  
Patent: WO 0105836-A 31 25-JAN-2001;  
Genentech, Inc. (US)  
Location/Qualifiers  
1. 1266  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

BASE COUNT 216 a 418 c 390 g 242 t  
ORIGIN

Query Match 1.9%; Score 32; DB 6; Length 1266;  
Best Local Similarity 100.0%; Pred. No. 1.2e-05;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 375 GTACCCCTGCTGCTGATGCTGCTGCTGCTG 406  
|||||  
Db 136 GTACCCCTGCTGCTGATGCTGCTGCTGCTG 167

RESULT 14  
AX464186 1266 bp DNA linear PAT 16-JUL-2002  
LOCUS Sequence 319 from Patent WO0140466.  
DEFINITION  
ACCESSION AX464186  
VERSION AX464186.1 GI:21899109  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
AUTHORS 1 Baker, K.P., Beresini, M., DeForge, L., Desnoyers, L., Filvaroff, E.,  
Gao, M.Q., Gerritsen, M.E., Goddard, A., Godowski, P.J., Gurney, A.L.,  
Sherwood, S., Smith, Y., Stewart, T.A., Tumas, D., Watanabe, C.K.,  
Wood, W.L. and Zhang, Z.  
TITLE Secretd and transmembrane polypeptides and nucleic acids encoding  
Same  
JOURNAL Patent: WO 0140466-A 319 07-JUN-2001;  
FEATURES Location/Qualifiers  
source 1.1266  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

BASE COUNT 216 a 418 c 390 g 242 t  
ORIGIN

Query Match 1.9%; Score 32; DB 6; Length 1266;  
Best Local Similarity 100.0%; Pred. No. 1.2e-05;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 375 GTACCCCTGCTGCTGATGCTGCTGCTGCTG 406  
|||||  
Db 136 GTACCCCTGCTGCTGATGCTGCTGCTGCTG 167

RESULT 15  
AF083500 1283 bp mRNA linear PRI 04-NOV-1998  
LOCUS Homo sapiens connective tissue growth factor-like protein  
DEFINITION  
ACCESSION AF083500  
VERSION AF083500.1 GI:3462835  
KEYWORDS  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
AUTHORS Kumar, S., Hand, A.T., Connor, J.C., Dodds, R.A., Ryan, P.J.,  
Trill, J.G., Fisher, S.M., Slemon, J.R., Lipschutz, D.B.,  
Barholomew, V., James, I.E., Rleman, D.J., Gowen, M. and Lee, J.C.

TITLE Identification and cloning of CTGF-L from human osteoblasts, a  
novel cysteine rich protein containing an IGF binding domain  
Bone 23 (5), S240 (1998)  
2 (bases 1 to 1283)

JOURNAL Kumar, S.  
REFERENCE Direct Submission  
AUTHORS Submitted (11-AUG-1998) Bone & Cartilage Biology, UW 2109,  
TITLE Smithkline Beecham, 709 Swedeland Road, King of Prussia, PA 19406,  
JOURNAL USA

FEATURES Location/Qualifiers  
source 1.1283  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="20"  
/map="20q12-q13"  
/cell\_type="primary osteoblast"  
9..761

CDS

/note="CTGF-L: encodes IGF binding (IGFB), Von Willebrand  
factor type C (VWC) and thrombospondin type I (TSP1)  
domains; member of the CCN (CTGF/Cyrr61/nov) family; lacks  
the fourth carboxy-terminal domain present in other  
members of the CCN family"  
/codon\_start=1  
/product="connective tissue growth factor-like protein  
precursor"  
/protein\_id="AACT0350.1"  
/db\_xref="GI:3462836"  
/translation="MRGTPKTHLAFSLICLLSKVRLQCPPTCPMPPPRCPGLVP  
LVLDGCGCRRVCARRLGEPCDQLHYCDASGLVCPGAGPGRGALCLLADDSCEV  
NGRLYREGFTFOPHCSTIRCEDEGFTCPICSDVRLPSMDCPHRREVLIGCCPE  
WVCGGGSLGTQPLPACGPGPSGLVSSLPVPCPEPSTANGPCSTGGLMATRVSN  
QNRCTLETQRRLLDLSRCPSPSGRSPNSAF"

BASE COUNT 235 a 418 c 389 g 241 t  
ORIGIN

Query Match 1.9%; Score 32; DB 9; Length 1283;  
Best Local Similarity 100.0%; Pred. No. 1.2e-05;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 375 GTACCCCTGCTGCTGATGCTGCTGCTGCTG 406  
|||||  
Db 135 GTACCCCTGCTGCTGATGCTGCTGCTGCTG 166

Search completed: July 28, 2003, 21:14:39  
Job time : 4755.55 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 14:31:11 ; Search time 448.586 Seconds  
(without alignments)  
8574.520 Million cell updates/sec

Title: US-10-010-408-1

Perfect score: 1708  
Sequence: 1 GACGCTTCGATCTCCAGAG.....GCCTAGAAATTAACACCCCAA 1708

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 2185239 seqs, 112599159 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N\_Geneseq\_101002.\*

```
1: /SID52/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID52/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1          | 1708  | 100.0       | 1708   | 20    | AAZ07516    |
| 2          | 753   | 44.1        | 753    | 20    | AAZ07517    |
| 3          | 681   | 39.9        | 681    | 20    | AAZ07521    |
| 4          | 354   | 20.7        | 439    | 24    | ABL59575    |
| 5          | 210   | 12.3        | 210    | 20    | AAZ07519    |
| 6          | 177   | 10.4        | 177    | 20    | AAZ07518    |
| 7          | 174   | 10.2        | 174    | 20    | AAZ07520    |
| 8          | 90    | 5.3         | 753    | 20    | AAZ07520    |
| 9          | 90    | 5.3         | 1734   | 20    | AAZ07520    |

|    |    |     |       |    |          |                    |
|----|----|-----|-------|----|----------|--------------------|
| 10 | 54 | 3.2 | 65    | 24 | ABN30189 | Rat spliced transc |
| 11 | 32 | 1.9 | 199   | 22 | ABA72245 | Human foetal liver |
| 12 | 32 | 1.9 | 199   | 22 | ABA38112 | Probe #16578 for g |
| 13 | 32 | 1.9 | 199   | 22 | AAK20657 | Human brain expres |
| 14 | 32 | 1.9 | 199   | 22 | AAK46811 | Human bone marrow  |
| 15 | 32 | 1.9 | 199   | 22 | AAI25749 | Probe #15682 for g |
| 16 | 32 | 1.9 | 199   | 22 | AAI52652 | Human genome-deriv |
| 17 | 32 | 1.9 | 199   | 24 | ABS21124 | Human foetal liver |
| 18 | 32 | 1.9 | 586   | 22 | ABA59703 | Probe #6598 for ge |
| 19 | 32 | 1.9 | 586   | 22 | ABA28232 | Human brain expres |
| 20 | 32 | 1.9 | 586   | 22 | AAK07972 | Human bone marrow  |
| 21 | 32 | 1.9 | 586   | 22 | AAK33842 | Probe #6485 for ge |
| 22 | 32 | 1.9 | 586   | 22 | AAI16552 | Probe #8254 used t |
| 23 | 32 | 1.9 | 586   | 22 | AAI39568 | Human genome-deriv |
| 24 | 32 | 1.9 | 586   | 24 | ABS08689 | Human genome-deriv |
| 25 | 32 | 1.9 | 738   | 20 | AAZ76501 | Human WISP-2 prote |
| 26 | 32 | 1.9 | 750   | 20 | AAZ76487 | Human WISP-2 prote |
| 27 | 32 | 1.9 | 841   | 20 | AAZ76502 | Human WISP-2 prote |
| 28 | 32 | 1.9 | 1257  | 20 | AAZ8435  | EGF-like homologue |
| 29 | 32 | 1.9 | 1266  | 21 | AAA30048 | Human PRO261 nucle |
| 30 | 32 | 1.9 | 1266  | 22 | AAZ21403 | Human CDNA sequenc |
| 31 | 32 | 1.9 | 1266  | 22 | AAZ60358 | PRO261 coding sequ |
| 32 | 32 | 1.9 | 1266  | 22 | AAZ97451 | Human angiotensin  |
| 33 | 32 | 1.9 | 1267  | 21 | AAZ75566 | Human PRO261 CDNA  |
| 34 | 32 | 1.9 | 1285  | 19 | AAZ29260 | Human connective t |
| 35 | 32 | 1.9 | 1293  | 20 | AAZ76486 | Human WISP-2 prote |
| 36 | 32 | 1.9 | 1309  | 22 | AAZ28214 | Connective tissue  |
| 37 | 32 | 1.9 | 1337  | 22 | AAH46952 | Human secreted pro |
| 38 | 32 | 1.9 | 1352  | 22 | AAH46936 | Human secreted pro |
| 39 | 32 | 1.9 | 1522  | 20 | AAH16595 | Human growth facto |
| 40 | 32 | 1.9 | 2136  | 22 | AAZ94706 | Human full-length  |
| 41 | 32 | 1.9 | 13255 | 22 | AAZ76535 | Human immune/haema |
| 42 | 27 | 1.6 | 51    | 20 | AAZ75559 | Human WISP-2 probe |
| 43 | 27 | 1.6 | 51    | 21 | AAZ75559 | Human PRO261 hybri |
| 44 | 27 | 1.6 | 51    | 21 | AAZ30051 | Probe for PRO261 n |
| 45 | 27 | 1.6 | 51    | 22 | AAZ60371 | PRO261 probe #1.   |

#### ALIGNMENTS

|          |   |
|----------|---|
| RESULT 1 |   |
| AAZ07516 |   |
| ID       | AAZ07516 standard; CDNA; 1708 BP.   |
| XX       | XX  |
| AC       | AAZ07516;   |
| XX       | XX  |
| DT       | 26-NOV-1999 (first entry)   |
| XX       | XX  |
| DE       | Rat HICP polypeptide encoding cDNA.   |
| XX       | XX  |
| KW       | Heparin-induced CCN-like protein; HICP; cell-associated activity; ss;   |
| KW       | cardiovascular disorder; aberrant cell proliferation; fibrotic disorder.  |
| OS       | Rattus sp.  |
| XX       | XX  |
| PN       | W09947556-A2.   |
| XX       | XX  |
| PD       | 23-SEP-1999.  |
| XX       | XX  |
| PF       | 18-MAR-1999; 99MO-US05999.  |
| XX       | XX  |
| PR       | 19-MAR-1998; 98US-0044273.  |
| XX       | XX  |
| PA       | (TUFT ) TUFTS COLLEGE.  |
| XX       | XX  |
| PI       | Castellot JJ;   |
| XX       | XX  |
| DR       | WPI: 1999-562060/47.  |
| XX       | XX  |
| PT       | P-PSDB: AAZ7434.  |
| XX       | XX  |
| PT       | Nucleic acid sequences encoding rat heparin-induced CCN-like protein, used in methods to identify modulators or in diagnostic applications. |

XX Claim 2: Fig 1: 108bp: English.

CC This cDNA encodes a rat heparin-induced CCN-like protein (HICP) protein.  
 CC Agents that stimulate or inhibit HICP protein activity or expression,  
 CC antisense HICP nucleic acid molecules and HICP antibodies, can be used to  
 CC modulate cell-associated activity. HICP modulators can be used to treat  
 CC disorders characterized by aberrant HICP protein activity or expression.  
 CC Probes capable of hybridizing to HICP mRNA or antibodies specific for  
 CC HICP can be used to detect HICP activity in a biological sample. HICP  
 CC can be used to treat disorders, such as a cardiovascular or fibrotic  
 CC disorder, characterized by aberrant cell proliferation.

XX Sequence 1708 BP; 362 A; 486 C; 478 G; 382 T; 0 other;

Query Match 100.0%; Score 1708; DB 20; Length 1708;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 1708; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGCTTCTGATCTCCAGAGACCCCTGGGTGGGACAGGAGCCCTTGGCAAGGCTGCAGCC 60  
 DB 1 GACGCTTCTGATCTCCAGAGACCCCTGGGTGGGACAGGAGCCCTTGGCAAGGCTGCAGCC 60  
 QY 61 GCTGGCACTGGCTTGGGATGAGGCTTATTACTGGGAACTGAGAGACTTAAGAGCTTC 120  
 DB 61 GCTGGCACTGGCTTGGGATGAGGCTTATTACTGGGAACTGAGAGACTTAAGAGCTTC 120  
 QY 121 CTGTGACGCTTCTTAAAGCTTGAAGCTTGTGGCTTGGGCTTGCACACTGTCAGA 180  
 DB 121 CTGTGACGCTTCTTAAAGCTTGAAGCTTGTGGCTTGGGCTTGCACACTGTCAGA 180  
 QY 181 CACCTTCTGTGTGGCTTCCAGGAGCTTCACTTCAAGTTTGAAGCTGTCCCAAGAGGAC 240  
 DB 181 CACCTTCTGTGTGGCTTCCAGGAGCTTCACTTCAAGTTTGAAGCTGTCCCAAGAGGAC 240  
 QY 241 ACGGTGACATGAGGGGACGCCACTGATCCATCTTGTGGCACTTCCCTCTGCTTC 300  
 DB 241 ACGGTGACATGAGGGGACGCCACTGATCCATCTTGTGGCACTTCCCTCTGCTTC 300  
 QY 301 TCTCAATGCTGTGTGCCAGCTGTGCGGACACCCCTGTAGCTGTCTTGGAGACACACCC 360  
 DB 301 TCTCAATGCTGTGTGCCAGCTGTGCGGACACCCCTGTAGCTGTCTTGGAGACACACCC 360  
 QY 361 AGTGGCCACAGGGGGATACCCCTGTGTGTGGATGCTGTGGCTGTCTTAAAGTGTGCAC 420  
 DB 361 AGTGGCCACAGGGGGATACCCCTGTGTGTGGATGCTGTGGCTGTCTTAAAGTGTGCAC 420  
 QY 421 GGAGGCTGGGGAGTCTGTGCGACACCTGATGTGTGGAGACCCGACGAGGCGTGGTT 480  
 DB 421 GGAGGCTGGGGAGTCTGTGCGACACCTGATGTGTGGAGACCCGACGAGGCGTGGTT 480  
 QY 481 GTGAGCCTGGGAGGAGCCCTGTGCGGACATGGGGCTGTGTCTCTTGGATGAGATGACG 540  
 DB 481 GTGAGCCTGGGAGGAGCCCTGTGCGGACATGGGGCTGTGTCTCTTGGATGAGATGACG 540  
 QY 541 GTAGCTGTAGGTGATGAGTGGCCGACAGTACCTGATGAGAGAGACCTTTAAACCAATGCA 600  
 DB 541 GTAGCTGTAGGTGATGAGTGGCCGACAGTACCTGATGAGAGAGACCTTTAAACCAATGCA 600  
 QY 601 GGGTCTGTGCGGCTGTGATGAGAGGAGTTCACCTGCGCTGGCGGCTGTCAAGTGAAGATG 660  
 DB 601 GGGTCTGTGCGGCTGTGATGAGAGGAGTTCACCTGCGCTGGCGGCTGTCAAGTGAAGATG 660  
 QY 661 TGCAGGCTGCCAGCTGGAGTGCACCGCCCAAGAGATACAGGTGCCAGGAAAGTCT 720  
 DB 661 TGCAGGCTGCCAGCTGGAGTGCACCGCCCAAGAGATACAGGTGCCAGGAAAGTCT 720  
 QY 721 GCCCGAGTGGGTATGTGACACAGGAGTACACCGGCGACAGGCGCTCCACGCGCAAG 780  
 DB 721 GCCCGAGTGGGTATGTGACACAGGAGTACACCGGCGACAGGCGCTCCACGCGCAAG 780  
 QY 781 GACACCAACTTCTGCGCTTGTACCTCGCTGCTGCTGCTGCTCTTCCAAATGGA 840  
 DB 781 GACACCAACTTCTGCGCTTGTACCTCGCTGCTGCTGCTGCTCTTCCAAATGGA 840

DB 781 GACACCAACTTCTGCGCTTGTACCTCGCTGCTGCTGCTGCTCTTCCAAATGGA 840  
 QY 841 GCACAGCCCTGGGGCCCTCTCAACACACCTGTGGGTGGGATGAGCAACCGAGTGTCA 900  
 DB 841 GCACAGCCCTGGGGCCCTCTCAACACACCTGTGGGTGGGATGAGCAACCGAGTGTCA 900  
 QY 901 ACCAGAACCGATTTGCGCACTGAGATGCAACGCGGCTGTGTGTGCGCAGACCTGCGC 960  
 DB 901 ACCAGAACCGATTTGCGCACTGAGATGCAACGCGGCTGTGTGTGCGCAGACCTGCGC 960  
 QY 961 TGGCAGCCAGGAGCCACACACTATGAGACAGTCTTTCTTAAGCCCACTGGGGATGCGGA 1020  
 DB 961 TGGCAGCCAGGAGCCACACACTATGAGACAGTCTTTCTTAAGCCCACTGGGGATGCGGA 1020  
 QY 1021 TACAGGCGCTGCGCATCTGAGCAATGACCTTACAGACAGGCGCTGACTGTGTAGAT 1080  
 DB 1021 TACAGGCGCTGCGCATCTGAGCAATGACCTTACAGACAGGCGCTGACTGTGTAGAT 1080  
 QY 1081 GCTCTTCTCCATGCTTGTGGCTGAGTAACTGTCTGCTGCTTGGATTCATGTTAGAGCC 1140  
 DB 1081 GCTCTTCTCCATGCTTGTGGCTGAGTAACTGTCTGCTGCTTGGATTCATGTTAGAGCC 1140  
 QY 1141 ACTGAGCGATCCCTGCTGTGTGAGTGAAGGAGGAGGAGTACACAGCTCCAGTTCTCTG 1200  
 DB 1141 ACTGAGCGATCCCTGCTGTGTGAGTGAAGGAGGAGGAGTACACAGCTCCAGTTCTCTG 1200  
 QY 1201 GTTACAGCCGGAATTCCTGGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260  
 DB 1201 GTTACAGCCGGAATTCCTGGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260  
 QY 1261 GACAAACCAAAAGACCTTTAAACCTAGAGCTATAGTGGCAACCTGGCCACCTGCTGG 1320  
 DB 1261 GACAAACCAAAAGACCTTTAAACCTAGAGCTATAGTGGCAACCTGGCCACCTGCTGG 1320  
 QY 1321 GATAGGCTCAATGTTAGGACGACAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 1380  
 DB 1321 GATAGGCTCAATGTTAGGACGACAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 1380  
 QY 1381 ACTTCTGTATGCTTGTCCCAAAAGATGATGAATCTGTAGTGTACCTTCCCTGACC 1440  
 DB 1381 ACTTCTGTATGCTTGTCCCAAAAGATGATGAATCTGTAGTGTACCTTCCCTGACC 1440  
 QY 1441 TGAGAACACCTTGTGCTGCGGAGATTCAGGGGAGCAATTTCTGTGAACATGAAGA 1500  
 DB 1441 TGAGAACACCTTGTGCTGCGGAGATTCAGGGGAGCAATTTCTGTGAACATGAAGA 1500  
 QY 1501 GATGATCAGACTGTGCTTAAAGAAATTCCTGAAGTCCAGAACTTGAAGCTTGTATT 1560  
 DB 1501 GATGATCAGACTGTGCTTAAAGAAATTCCTGAAGTCCAGAACTTGAAGCTTGTATT 1560  
 QY 1561 CAGGAATGACATCTTAAAGCAGTCCGCAAAACAGGAAGCTCAGACCTTGGCAGGCC 1620  
 DB 1561 CAGGAATGACATCTTAAAGCAGTCCGCAAAACAGGAAGCTCAGACCTTGGCAGGCC 1620  
 QY 1621 AGGGCTTTCTTCAAGATGAGAAAGACAGGAGACAGAGTACTCTCTCTGGAGA 1680  
 DB 1621 AGGGCTTTCTTCAAGATGAGAAAGACAGGAGACAGAGTACTCTCTCTGGAGA 1680  
 QY 1681 CTAGTCTAGCCTAGATAAACAACCCAAA 1708  
 DB 1681 CTAGTCTAGCCTAGATAAACAACCCAAA 1708

# RESULT 2

AA207517 standard: cDNA: 753 BP.

AA207517;

26-NOV-1999 (first entry)

Rat HICP polypeptide coding sequence.





Matches 681; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 318 CAGCTGTGCGGACACCTGTACTGTCTCTGGACACACCCAGTGGCCACAGGGGTA 377
   |||||||
Db 1 CAGCTGTGCGGACACCTGTACTGTCTCTGGACACACCCAGTGGCCACAGGGGTA 60
OY 378 CCCCTGTGCTGGATGGCTGTGGCTGTGTAAAGTGTGTGACGAGAGCTGGGGGATCC 437
   |||||||
Db 61 CCCCTGTGCTGGATGGCTGTGGCTGTGTAAAGTGTGTGACGAGAGCTGGGGGATCC 120
OY 438 TCGGACACCTGATGTGTGACACCCACGAGGCGGCTGTTCAGCTCTGGGGAGGC 497
   |||||||
Db 121 TCGGACACCTGATGTGTGACACCCACGAGGCGGCTGTTCAGCTCTGGGGAGGC 180
OY 498 CCTGGGCGCATGGGGCTGTGTCTCTTGATGAGATGAGCTAGCTGTAGGTAT 557
   |||||||
Db 181 CCTGGGCGCATGGGGCTGTGTCTCTTGATGAGATGAGCTAGCTGTAGGTAT 240
OY 558 GCGCCGAGTACTGTGATGAGAGACCTTTAAACCAATGACAGGGTCTGTGCCGTGT 617
   |||||||
Db 241 GCGCCGAGTACTGTGATGAGAGACCTTTAAACCAATGACAGGGTCTGTGCCGTGT 300
OY 618 GATGACGGTGGCTTCACTGCGCGCTGTGAGTGTGAGTGTGGCGGCTGGCCAGCTGG 677
   |||||||
Db 301 GATGACGGTGGCTTCACTGCGCGCTGTGAGTGTGAGTGTGGCGGCTGGCCAGCTGG 360
OY 678 GACTGCCACGCGCCCAAGAAATACAGTGTGCCAGAAAGTGTGCCCGAGTGGTATGT 737
   |||||||
Db 361 GACTGCCACGCGCCCAAGAAATACAGTGTGCCAGAAAGTGTGCCCGAGTGGTATGT 420
OY 738 GACGAGGAGTGTACACCGCGGATCCAGGCTCCAGGGCGCAAGGACCACTTTCTGCC 797
   |||||||
Db 421 GACGAGGAGTGTACACCGCGGATCCAGGCTCCAGGGCGCAAGGACCACTTTCTGCC 480
OY 798 CTTTGATCACTGCTCTGCTGTATGCTCTTGTCCAAATGTGAGACACAGCTGGGGCCCC 857
   |||||||
Db 481 CTTTGATCACTGCTCTGCTGTATGCTCTTGTCCAAATGTGAGACACAGCTGGGGCCCC 540
OY 858 TGCCTAACCACTGTGGCTGGGCAATAGCCACCCGAGTGTCCAAACCAAGCAATTTCTGC 917
   |||||||
Db 541 TGCCTAACCACTGTGGCTGGGCAATAGCCACCCGAGTGTCCAAACCAAGCAATTTCTGC 600
OY 918 CAACGTGGAGATCCAGCGCGCTGTGTCTGCCAGACCCCTGGCTGGACAGGAGCCAC 977
   |||||||
Db 601 CAACGTGGAGATCCAGCGCGCTGTGTCTGCCAGACCCCTGGCTGGACAGGAGCCAC 660
OY 978 AGCTCATGAGACAGTGTCTTC 998
   |||||||
Db 661 AGCTCATGAGACAGTGTCTTC 681

```

RESULT 4  
ABLS9575  
ID ABL59575 standard; DNA: 439 BP.  
XX ABL59575;  
AC  
XX  
XX  
XX 17-JUL-2002 (first entry)  
XX  
XX  
XX Rat OST23 gene fragment SEQ ID NO:23.  
XX  
XX Rat; OST: osteoregenerative; parathyroid hormone; tibiae; osteopathic;  
XX gene therapy; parathyroid hormone receptor ligand; bone disorder;  
XX bone formation disorder; bone resorption disorder; osteoporosis;  
XX osteopenia; osteopetrosis; gene: ds.  
OS  
XX Ratus sp.  
PN WO200224943-A2.  
XX  
XX 28-MAR-2002.  
XX  
XX 19-SEP-2001; 2001WO-US29548.  
PF

XX 19-SEP-2000; 2000US-233579P.  
XX  
XX (CURA-) CURAGEN CORP.  
XX (GLAXO) GLAXO GROUP LTD.  
XX  
XX  
XX Horsesovsky GJ, Noel LS, Raha D;  
XX WPI; 2002-401989/43.  
XX  
XX  
XX Identifying parathyroid hormone receptor ligands and osteoregenerative  
XX agents involves detecting the expression of nucleic acids which are  
XX regulated by parathyroid hormone  
XX  
XX  
XX Claim 45; Page 22; 90pp; English.

The present invention describes a method (M1) for identifying parathyroid hormone receptor ligands (I) and osteoregenerative agents by contacting a test cell population (CP) comprising cells expressing nucleic acid sequences (S) of OST-47 and 48, with a test agent, measuring nucleic acid sequence expression, comparing it with reference CP and identifying the ligand and agent by the difference in expression levels. Also described is a method (M2) for treating a bone disorder in a subject by administering to the subject an agent that modulates the expression or activity of (S). (I) have osteopathic activities, and can be used to modulate the expression of OST 1-48, and can also be used in gene therapy. (M1) is useful for identifying parathyroid hormone receptor ligands (I) or osteoregenerative agents. OST gene sequence can be used for assessing the osteoregenerative activity of a test agent in a subject, and for diagnosing or determining the susceptibility to bone disorder and assessing efficacy of a treatment of a bone disorder in a subject, human or rodent. (M2) is useful for treating a bone disorder including bone formation disorder or bone resorption disorder (e.g. osteoporosis, osteopenia and osteopetrosis). OST polypeptides are useful as immunogens to raise anti-OST antibodies (II). Parathyroid hormone receptor ligands (I) identified by (M1) are useful for treating bone disorders. The present sequence represents an OST23 gene fragment isolated from rat tibiae, from the present invention.

Sequence 439 BP; 128 A; 108 C; 99 G; 104 T; 0 other;

Query Match 20.7%; Score 354; DB 24; Length 439;

Best Local Similarity 99.8%; Pred. No. 4e-162; Mismatches 1; Indels 0; Gaps 0;

Matches 404; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

OY 1250 TGTACAAAAGGACACCAAAAGACCTTTAACTAGGCTTACTGGCAAACTGGCC 1309
   |||||||
Db 1 TGTACAAAAGGACACCAAAAGACCTTTAACTAGGCTTACTGGCAAACTGGCC 60
OY 1310 ACCGTGTGGGGATAGTCAATGTTAGACACGACAGCAGATTCCTGAAACTTCAT 1369
   |||||||
Db 61 ACCGTGTGGGGATAGTCAATGTTAGACACGACAGCAGATTCCTGAAACTTCAT 120
OY 1370 TCCCTTTGACCTTCTGATCTGTCTGCCAAAGATGTGAATGAACTGTAAGTATAC 1429
   |||||||
Db 121 TCCCTTTGACCTTCTGATCTGTCTGCCAAAGATGTGAATGAACTGTAAGTATAC 180
OY 1430 CTTCCCTGACCTGAGAACACCTGCTGTGCGGAAGTATAGGGGCGAATTTCTGT 1489
   |||||||
Db 181 CTTCCCTGACCTGAGAACACCTGCTGTGCGGAAGTATAGGGGCGAATTTCTGT 240
OY 1490 GAACATGAAGAGATGATACACTGTCTTAAAGAAATTCCTGAAGTCCAGAACTTGAG 1549
   |||||||
Db 241 GAACATGAAGAGATGATACACTGTCTTAAAGAAATTCCTGAAGTCCAGAACTTGAG 300
OY 1550 CTTTGATTTTTCAGGAATGACATCTTTAAGCACTGCGAAACAGGAAGCTCCACACC 1609
   |||||||
Db 301 CTTTGATTTTTCAGGAATGACATCTTTAAGCACTGCGAAACAGGAAGCTCCACACC 360
OY 1610 TCTGGAGGCGCAGGGGCTTCTCTTACAGATGAGAAAGCAAGGG 1654
   |||||||
Db 361 TCTGGAGGCGCAGGGGCTTCTCTTACAGATGAGAAAGCAAGGG 405

```





XX Wnt-1 induced secreted protein: WISP-1; WISP-2; WISP-3; CTGF; tumour;  
KM connective tissue growth factor; cancer; melanoma; arteriosclerosis;  
KM leukaemia; lymphoid malignancy; haematopoiesis-related disorder;  
KM tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;  
KM kidney disorder; bone-related disorder; osteoporosis; trauma; burn;  
KM connective tissue disorder; catabolic state; inflammation;  
KM testicular-related disorder; angiogenesis; immunological disorder; ss.  
XX  
XX Mus sp.  
XX  
XX W09921998-A1.  
XX  
XX 06-MAY-1999.  
XX  
XX 29-OCT-1998; 98WO-US22991.  
XX  
XX 14-APR-1998; 98US-0081695.  
XX 29-OCT-1997; 97US-0063704.  
XX 03-FEB-1998; 98US-0073612.  
XX  
XX (GETH ) GENENTECH INC.  
XX Botstein DA, Cohen RL, Goddard A, Gurney AL, Hillan K;  
PI Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WI;  
XX  
XX WPI: 1999-337420/28.  
XX P-PSDB: AAY17651.  
XX  
XX New Isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3  
XX  
XX Example 2: Page 178-179; 284pp; English.  
XX  
XX The present invention describes Wnt-1 induced secreted polypeptides,  
CC WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2  
CC and WISP-3 have homology to connective tissue growth factor (CTGF).  
CC Products from the present invention can be used to treat WISP-related  
CC disorders such as breast, ovarian, and colon cancer or melanoma. The  
CC products can be used to treat arteriosclerosis. The products can also be  
CC used to treat other diseases e.g. benign and malignant tumours,  
CC leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,  
CC hypothalamic and other glandular, macrophagal, epithelial, stromal, and  
CC blastocoeleic disorders, haematoplasia-related disorders, tissue-growth  
CC disorders, skin disorders, desmoplasia, fibrotic lesions, kidney  
CC disorders, bone-related disorders such as osteoporosis, trauma such as  
CC burns, incisions, and other wounds, connective tissue disorders,  
CC catabolic states, testicular-related disorders, and inflammatory,  
CC angiogenic and immunologic disorders including arteriosclerosis. The  
CC products can also be used for detection and diagnosis especially of  
CC individuals with neoplastic cell growth or proliferation. The products  
CC can be used in the production of transgenic or knock-out animals.  
CC Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing  
CC cells.  
XX  
XX S0 Sequence 1734 BP; 355 A; 491 C; 495 G; 393 T; 0 other;  
XX  
XX Query Match 5.3%; Score 90; DB 20; Length 1734;  
XX Best Local Similarity 100.0%; Pred. No. 1.9e-33;  
XX Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

AC ABN30189;  
XX  
XX 15-JUL-2002 (first entry)  
XX  
XX Rat spliced transcript detection oligonucleotide SEQ ID NO:2937.  
DE  
XX Human; mouse; rat; splice transcript; detection; RNA transcript;  
XX splice variant; transcriptome; oligonucleotide library; ss.  
XX  
XX Rattus norvegicus.  
XX  
XX W0200210449-A2.  
XX  
XX 07-FEB-2002.  
XX  
XX 20-JUL-2001; 2001WO-IB01903.  
XX  
XX 28-JUL-2000; 2000US-221607P.  
XX 02-MAY-2001; 2001US-287724P.  
XX  
XX (COMP-) COMPUGEN INC.  
XX  
XX Shoshan A, Wasserman A, Mintz E, Mintz L, Falgler S;  
PI WPI: 2002-257383/30.  
XX  
XX  
XX New oligonucleotide libraries comprising oligonucleotides which  
XX selectively hybridize to mRNAs transcribed from a transcription unit of  
XX a genome, useful for detecting tissue-, pathology-, and  
XX developmental-specific genes  
XX  
XX Example 1: SEQ ID 2937; 47pp; English.  
XX  
XX The present invention describes oligonucleotide libraries for detecting  
XX messenger RNAs that populate a (sub-)transcriptome, where the  
XX (sub-)transcriptome comprises messenger RNAs transcribed from multiple  
XX transcription units that populate a genome. The library comprises  
XX several oligonucleotides, each capable of hybridizing selectively to a  
XX set of messenger RNAs transcribed from a given transcription unit of  
XX the genome, which encodes one or more messenger RNA splice variants.  
XX The oligonucleotide libraries are useful for detecting mRNAs from a  
XX biological sample, in expression profiling studies, in qualitatively or  
XX quantitatively characterizing the corresponding transcriptome, and in  
XX detecting RNA transcripts and splice variants of human or animal  
XX transcriptomes. The libraries may also be used as specialized mini  
XX libraries to detect transcripts of a sub-transcriptome under a  
XX particular biological or pathological state, and so allowing the  
XX detection of tissue- and pathology-specific genes such as those genes  
XX only expressed in specific tissue under a specific pathological  
XX condition; to detect developmental specific genes; and to detect RNA  
XX transcripts and splice variants of a transcriptome of a patient suffering  
XX from a particular disorder. ABN27253 to ABN59589 represent  
XX oligonucleotide sequences from rats, humans and mice, which are used in  
XX the exemplification of the present invention.  
XX N.B. The sequence data for this patent did not form part of the printed  
XX specification, but was obtained in electronic format directly from WIPO  
XX at ftp.wipo.int/pub/published\_pcl\_sequences.  
XX  
XX S0 Sequence 65 BP; 21 A; 20 C; 13 G; 11 T; 0 other;  
XX  
XX Query Match 3.2%; Score 54; DB 24; Length 65;  
XX Best Local Similarity 100.0%; Pred. No. 6.7e-16;  
XX Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 10  
ABN30189  
ID ABN30189 standard; DNA: 65 BP.  
XX

RESULT 11  
ABA72245  
ID ABA72245 standard; DNA: 199 BP.  
XX

AC ABA72245;  
 XX  
 PD  
 DT 01-FEB-2002 (first entry)  
 XX  
 XX Human foetal liver single exon nucleic acid probe #20550.  
 DE  
 XX Human; foetal liver; gene expression; single exon nucleic acid probe; ss.  
 KW  
 XX Homo sapiens.  
 OS  
 XX WO200157277-A2.  
 PN  
 XX 09-AUG-2001.  
 PD  
 XX  
 PF 30-JAN-2001; 2001WO-US00669.  
 XX  
 XX 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 PA  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 XX WPI; 2001-483447/52.  
 DR  
 XX Human genome-derived single exon nucleic acid probes useful for  
 PT analyzing gene expression in human foetal liver.  
 PS  
 XX Claim 4; SEQ ID NO 20550; 639pp + sequence listing; English.  
 XX  
 CC The invention relates to a single exon nucleic acid probe for  
 CC measuring human gene expression in a sample derived from human foetal  
 CC liver. The single exon nucleic acid probes may be used for predicting,  
 CC measuring and displaying gene expression in samples derived from human  
 CC foetal liver. The present sequence is a single exon nucleic acid  
 CC probe of the invention.  
 CC Note: The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pcl\_sequences.  
 CC  
 XX Sequence 199 BP; 32 A; 65 C; 73 G; 29 T; 0 other;  
 SQ  
 Query Match 1.9%; Score 32; DB 22; Length 199;  
 Best Local Similarity 100.0%; Pred. No. 3.6e-05;  
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 654 GAGGATGTCGGCTGCCAGCTGGGACTGCC 685  
 DB 129 GAGGATGTCGGCTGCCAGCTGGGACTGCC 160  
 RESULT 12  
 ABA38112  
 ID ABA38112 standard; DNA; 199 BP.  
 XX  
 AC ABA38112;  
 XX  
 XX 23-JAN-2002 (first entry)  
 DT  
 XX  
 XX Probe #16578 for gene expression analysis in human heart cell sample.  
 DE  
 XX Human; gene expression; heart; microarray; vascular system; probe;  
 KW cardiovascular disease; hypertension; cardiac arrhythmia;  
 KW congenital heart disease; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157274-A2.

XX  
 XX 09-AUG-2001.  
 PD  
 DT 30-JAN-2001; 2001WO-US00666.  
 XX  
 XX  
 PF 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 PA  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 XX WPI; 2001-48899/53.  
 DR  
 XX Single exon nucleic acid probes for analyzing gene expression in human  
 PT hearts -  
 PS  
 XX Claim 4; SEQ ID NO 16578; 530pp; English.  
 XX  
 CC The present invention relates to single exon nucleic acid probes for  
 CC measuring human gene expression in a sample derived from human heart. The  
 CC present sequence is one such probe. The probes may be used for  
 CC predicting, measuring and displaying gene expression in samples derived  
 CC from the human heart via microarrays. By measuring gene expression, the  
 CC probes are useful for predicting, diagnosing, grading, staging,  
 CC monitoring and prognosing diseases of the human heart and vascular system  
 CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and  
 CC congenital heart disease.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pcl\_sequences.  
 CC  
 XX Sequence 199 BP; 32 A; 65 C; 73 G; 29 T; 0 other;  
 SQ  
 Query Match 1.9%; Score 32; DB 22; Length 199;  
 Best Local Similarity 100.0%; Pred. No. 3.6e-05;  
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 654 GAGGATGTCGGCTGCCAGCTGGGACTGCC 685  
 DB 129 GAGGATGTCGGCTGCCAGCTGGGACTGCC 160  
 RESULT 13  
 AAK20667  
 ID AAK20667 standard; DNA; 199 BP.  
 XX  
 AC AAK20667;  
 XX  
 XX 05-NOV-2001 (first entry)  
 DT  
 XX  
 XX Human brain expressed single exon probe SEQ ID NO: 20658.  
 DE  
 XX Human; brain expressed exon; gene expression analysis; probe;  
 KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;  
 KW epilepsy; cancer; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157275-A2.  
 PD  
 XX 09-AUG-2001.  
 DT  
 XX 30-JAN-2001; 2001WO-US00667.  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.



Query Match 1.9%; Score 32; DB 22; Length 199;  
 Best Local Similarity 100.0%; Pred. No. 3.6e-05;  
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 654 GAGGATGTGGGGTGGCCAGCTGGGACTGGCC 685  
 ||||||||||||||||||||||||||||||||  
 Db 129 GAGGATGTGGGGTGGCCAGCTGGGACTGGCC 160

Search completed: July 28, 2003, 15:54:55  
 Job time : 452.586 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 14:50:06 ; Search time 93.3356 Seconds  
(without alignments)  
5612.050 Million cell updates/sec

Title: US-10-010-408-1

Perfect score: 1708

Sequence: 1 GACGCTTCGATCTCCAGAG.....GCCTAGATTAACACCCAAA 1708

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 441362 seqs, 15338381 residues

Word size : 0

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued\_Patents\_NA:\*  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PTUS.COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID                | Description       |
|------------|-------|-------------|--------|----------------------|-------------------|
| 1          | 90    | 5.3         | 1734   | 4 US-09-182-145-17   | Sequence 17, Appl |
| 2          | 90    | 5.3         | 1734   | 4 US-09-182-145-18   | Sequence 18, Appl |
| 3          | 32    | 1.9         | 738    | 4 US-09-182-145-38   | Sequence 38, Appl |
| 4          | 32    | 1.9         | 841    | 4 US-09-182-145-39   | Sequence 39, Appl |
| 5          | 32    | 1.9         | 1293   | 4 US-09-182-145-13   | Sequence 13, Appl |
| 6          | 32    | 1.9         | 1293   | 4 US-09-182-145-14   | Sequence 14, Appl |
| 7          | 27    | 1.6         | 31     | 4 US-09-182-145-117  | Sequence 117, App |
| 8          | 19    | 1.1         | 616    | 4 US-09-385-982-220  | Sequence 220, App |
| 9          | 19    | 1.1         | 1196   | 4 US-09-149-476-225  | Sequence 225, App |
| 10         | 19    | 1.1         | 1220   | 4 US-09-149-476-57   | Sequence 57, Appl |
| 11         | 19    | 1.1         | 1514   | 2 US-09-213-768-1    | Sequence 1, Appl  |
| 12         | 19    | 1.1         | 1539   | 4 US-09-668-680-13   | Sequence 13, Appl |
| 13         | 18    | 1.1         | 20     | 2 US-09-213-768-2    | Sequence 2, Appl  |
| 14         | 18    | 1.1         | 2104   | 4 US-09-313-930-1    | Sequence 1, Appl  |
| 15         | 18    | 1.1         | 2949   | 4 US-08-259-451-10   | Sequence 10, Appl |
| 16         | 18    | 1.1         | 3727   | 1 US-08-249-380-1    | Sequence 1, Appl  |
| 17         | 18    | 1.1         | 8957   | 4 US-08-259-451-1    | Sequence 1, Appl  |
| 18         | 17    | 1.0         | 24     | 4 US-09-182-145-110  | Sequence 110, App |
| 19         | 17    | 1.0         | 44     | 4 US-09-182-145-152  | Sequence 152, App |
| 20         | 17    | 1.0         | 454    | 4 US-08-691-814B-117 | Sequence 117, App |
| 21         | 17    | 1.0         | 480    | 3 US-09-188-930-206  | Sequence 206, App |
| 22         | 17    | 1.0         | 482    | 2 US-08-691-814B-120 | Sequence 120, App |
| 23         | 17    | 1.0         | 614    | 4 US-08-998-416-151  | Sequence 151, App |
| 24         | 17    | 1.0         | 661    | 4 US-08-578-634C-3   | Sequence 3, Appl  |
| 25         | 17    | 1.0         | 661    | 4 US-09-430-010-3    | Sequence 3, Appl  |
| 26         | 17    | 1.0         | 742    | 2 US-08-966-316-2    | Sequence 2, Appl  |
| 27         | 17    | 1.0         | 896    | 3 US-09-188-930-36   | Sequence 36, Appl |

|    |    |     |      |                     |                   |
|----|----|-----|------|---------------------|-------------------|
| 28 | 17 | 1.0 | 933  | 3 US-08-987-743-1   | Sequence 1, Appl  |
| 29 | 17 | 1.0 | 1308 | 3 US-08-987-743-5   | Sequence 5, Appl  |
| 30 | 17 | 1.0 | 1491 | 4 US-09-662-249A-3  | Sequence 3, Appl  |
| 31 | 17 | 1.0 | 1743 | 3 US-08-665-259-20  | Sequence 20, Appl |
| 32 | 17 | 1.0 | 1743 | 3 US-08-762-500-20  | Sequence 20, Appl |
| 33 | 17 | 1.0 | 1821 | 4 US-09-149-476-90  | Sequence 90, Appl |
| 34 | 17 | 1.0 | 1974 | 3 US-08-762-500-78  | Sequence 78, Appl |
| 35 | 17 | 1.0 | 2329 | 4 US-08-455-559-9   | Sequence 9, Appl  |
| 36 | 17 | 1.0 | 2329 | 4 US-09-145-060-9   | Sequence 9, Appl  |
| 37 | 17 | 1.0 | 2329 | 5 PCT-US94-00657-9  | Sequence 9, Appl  |
| 38 | 17 | 1.0 | 2517 | 3 US-08-733-360A-4  | Sequence 4, Appl  |
| 39 | 17 | 1.0 | 2517 | 3 US-08-987-743-16  | Sequence 16, Appl |
| 40 | 17 | 1.0 | 2517 | 4 US-08-916-935-4   | Sequence 4, Appl  |
| 41 | 17 | 1.0 | 3280 | 4 US-09-484-970B-43 | Sequence 43, Appl |
| 42 | 17 | 1.0 | 3865 | 4 US-09-149-476-296 | Sequence 296, App |
| 43 | 17 | 1.0 | 6803 | 3 US-08-665-259-19  | Sequence 19, Appl |
| 44 | 17 | 1.0 | 6803 | 3 US-08-762-500-19  | Sequence 19, Appl |
| 45 | 17 | 1.0 | 7874 | 4 US-09-780-175-96  | Sequence 96, Appl |

## ALIGNMENTS

```

RESULT 1
US-09-182-145-17
: Sequence 17, Application US/09182145B
: Patent No. 6387657
: GENERAL INFORMATION:
: APPLICANT: Botstein, David A.
: APPLICANT: Cohen, Robert
: APPLICANT: Goddard, Audrey
: APPLICANT: Gurney, Austin L.
: APPLICANT: Hillan, Kenneth J.
: APPLICANT: Lawrence, David A.
: APPLICANT: Levine, Arnold J.
: APPLICANT: Pennica, Diane
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Wood, William I.
: TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
: FILE REFERENCE: P117682
: CURRENT APPLICATION NUMBER: US/09/182.145B
: EARLIER FILING DATE: 1998-10-29
: EARLIER APPLICATION NUMBER: US 60/063,704
: EARLIER FILING DATE: 1997-10-29
: EARLIER APPLICATION NUMBER: US 60/073,612
: EARLIER FILING DATE: 1998-02-04
: EARLIER APPLICATION NUMBER: US 60/081,695
: EARLIER FILING DATE: 1998-04-14
: NUMBER OF SEQ ID NOS: 156
: SEQ ID NO 17
: LENGTH: 1734
: TYPE: DNA
: ORGANISM: Mus musculus
: US-09-182-145-17

Query Match          5.3%: Score 90; DB 4; Length 1734;
Best Local Similarity 100.0%: Pred. No. 1.9e-34;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 410 AGTGTGACGAGAGGCTGGGGAGTCTCGACACCTGCATGTCGACCCACCA 469
    |||||||
Db 418 AGTGTGACGAGAGGCTGGGGAGTCTCGACACCTGCATGTCGACCCACCA 477
    |||||||

QY 470 GGGCCTGTTTGTGACGCTGGGGAGGCC 499
    |||||||
Db 478 GGGCCTGTTTGTGACGCTGGGGAGGCC 507
    |||||||

RESULT 2
US-09-182-145-18/c
: Sequence 18, Application US/09182145B
: Patent No. 6387657
: GENERAL INFORMATION:

```

APPLICANT: Botstein, David A.  
APPLICANT: Cohen, Robert  
APPLICANT: Goddard, Audrey  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: Lawrence, David A.  
APPLICANT: Levine, Arnold J.  
APPLICANT: Pennica, Diane  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME  
FILE REFERENCE: P1176R2  
CURRENT APPLICATION NUMBER: US/09/182,145B  
EARLIER FILING DATE: 1998-10-29  
EARLIER APPLICATION NUMBER: US 60/063,704  
EARLIER FILING DATE: 1997-10-29  
EARLIER APPLICATION NUMBER: US 60/073,612  
EARLIER FILING DATE: 1998-02-04  
EARLIER APPLICATION NUMBER: US 60/081,695  
EARLIER FILING DATE: 1998-04-14  
NUMBER OF SEQ ID NOS: 156  
SEQ ID NO 18  
LENGTH: 1734  
TYPE: DNA  
ORGANISM: Mus musculus  
US-09-182-145-18

Query Match 5.3%: Score 90; DB 4; Length 1734;  
Best Local Similarity 100.0%; Pred. No. 1.9e-31;  
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 410 AGTGTGTGCACGAGAGCTGGGGAGTCTCTGCACACGCGATGTCTGGACCCAGCCA 469  
DB 1317 AGTGTGTGCACGAGAGCTGGGGAGTCTCTGCACACGCGATGTCTGGACCCAGCCA 1258  
QY 470 GGGCCTGGTTTCTCAGCCCTGGGGCAGGCC 499  
DB 1257 GGGCCTGGTTTCTCAGCCCTGGGGCAGGCC 1228

RESULT 3  
US-09-182-145-38  
Sequence 38, Application US/09182145B  
Patent No. 6387657

GENERAL INFORMATION:  
APPLICANT: Botstein, David A.  
APPLICANT: Cohen, Robert  
APPLICANT: Goddard, Audrey  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: Lawrence, David A.  
APPLICANT: Levine, Arnold J.  
APPLICANT: Pennica, Diane  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME  
FILE REFERENCE: P1176R2  
CURRENT APPLICATION NUMBER: US/09/182,145B  
EARLIER FILING DATE: 1998-10-29  
EARLIER APPLICATION NUMBER: US 60/063,704  
EARLIER FILING DATE: 1997-10-29  
EARLIER APPLICATION NUMBER: US 60/073,612  
EARLIER FILING DATE: 1998-02-04  
EARLIER APPLICATION NUMBER: US 60/081,695  
EARLIER FILING DATE: 1998-04-14  
NUMBER OF SEQ ID NOS: 156  
SEQ ID NO 38  
LENGTH: 738  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-182-145-38

Query Match 1.9%: Score 32; DB 4; Length 738;

Best Local Similarity 100.0%; Pred. No. 4.4e-06;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 375 GTACCCCTGCTGCTGATGCGCTGGCTCTG 406  
DB 115 GTACCCCTGCTGCTGATGCGCTGGCTCTG 146

RESULT 4  
US-09-182-145-39

Sequence 39, Application US/09182145B  
Patent No. 6387657  
GENERAL INFORMATION:  
APPLICANT: Botstein, David A.  
APPLICANT: Cohen, Robert  
APPLICANT: Goddard, Audrey  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: Lawrence, David A.  
APPLICANT: Levine, Arnold J.  
APPLICANT: Pennica, Diane  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME  
FILE REFERENCE: P1176R2  
CURRENT APPLICATION NUMBER: US/09/182,145B  
EARLIER FILING DATE: 1998-10-29  
EARLIER APPLICATION NUMBER: US 60/063,704  
EARLIER FILING DATE: 1997-10-29  
EARLIER APPLICATION NUMBER: US 60/073,612  
EARLIER FILING DATE: 1998-02-04  
EARLIER APPLICATION NUMBER: US 60/081,695  
EARLIER FILING DATE: 1998-04-14  
NUMBER OF SEQ ID NOS: 156  
SEQ ID NO 39  
LENGTH: 841  
TYPE: DNA  
ORGANISM: Artificial sequence  
NAME/KEY: misc.feature  
LOCATION: 1-841  
OTHER INFORMATION: Sequence is synthesized.  
Patent No. 6387657  
US-09-182-145-39

Query Match 1.9%: Score 32; DB 4; Length 841;  
Best Local Similarity 100.0%; Pred. No. 4.4e-06;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 654 GAGATGTGCGGCTGCCAGCTGGGACTGCC 685  
DB 417 GAGATGTGCGGCTGCCAGCTGGGACTGCC 448

RESULT 5  
US-09-182-145-13

Sequence 13, Application US/09182145B  
Patent No. 6387657  
GENERAL INFORMATION:  
APPLICANT: Botstein, David A.  
APPLICANT: Cohen, Robert  
APPLICANT: Goddard, Audrey  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: Lawrence, David A.  
APPLICANT: Levine, Arnold J.  
APPLICANT: Pennica, Diane  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME  
FILE REFERENCE: P1176R2  
CURRENT APPLICATION NUMBER: US/09/182,145B  
EARLIER FILING DATE: 1998-10-29

```
; EARLIER APPLICATION NUMBER: US 60/063,704
; EARLIER FILING DATE: 1997-10-29
; EARLIER APPLICATION NUMBER: US 60/073,612
; EARLIER FILING DATE: 1998-02-04
; EARLIER APPLICATION NUMBER: US 60/081,695
; EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 13
; LENGTH: 1293
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-182-145-13
```

```
Query Match          1.9%; Score 32; DB 4; Length 1293;
Best Local Similarity 100.0%; Pred. No. 4.4e-06;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 375 GTACCCCTGTGCTGTGATGGCTGTGCTGCTG 406
Db 148 GTACCCCTGTGCTGTGATGGCTGTGCTGCTG 179
```

```
RESULT 6
US-09-182-145-14/c
; Sequence 14, Application US/09182145B
; Patent No. 6387657
; GENERAL INFORMATION:
```

```
; APPLICANT: Bolstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Kenneth L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/09/182,145B
; EARLIER APPLICATION NUMBER: US 60/063,704
; EARLIER FILING DATE: 1997-10-29
; EARLIER APPLICATION NUMBER: US 60/073,612
; EARLIER FILING DATE: 1998-02-04
; EARLIER APPLICATION NUMBER: US 60/081,695
; EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 14
; LENGTH: 1293
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-182-145-14
```

```
Query Match          1.9%; Score 32; DB 4; Length 1293;
Best Local Similarity 100.0%; Pred. No. 4.4e-06;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 375 GTACCCCTGTGCTGTGATGGCTGTGCTGCTG 406
Db 1146 GTACCCCTGTGCTGTGATGGCTGTGCTGCTG 1115
```

```
RESULT 7
US-09-182-145-117
```

```
; Sequence 117, Application US/09182145B
; Patent No. 6387657
; GENERAL INFORMATION:
```

```
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P1176R2
; CURRENT APPLICATION NUMBER: US/09/182,145B
; EARLIER APPLICATION NUMBER: US 60/063,704
; EARLIER FILING DATE: 1997-10-29
; EARLIER APPLICATION NUMBER: US 60/073,612
; EARLIER FILING DATE: 1998-02-04
; EARLIER APPLICATION NUMBER: US 60/081,695
; EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 117
; LENGTH: 51
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1-51
; OTHER INFORMATION: Sequence is synthesized.
US-09-182-145-117

Query Match          1.6%; Score 27; DB 4; Length 51;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 380 CCTGTGCTGTGATGGCTGTGCTGCTG 406
Db 1 CCTGTGCTGTGATGGCTGTGCTGCTG 27
```

```
RESULT 8
US-09-385-982-220/c
; Sequence 220, Application US/09385982
; Patent No. 6262334
; GENERAL INFORMATION:
```

```
; APPLICANT: ENDEGE, WILSON O., ET AL.
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; FILE REFERENCE: CCDDNA-260XX
; CURRENT APPLICATION NUMBER: US/09/385,982
; EARLIER FILING DATE: 1999-08-30
; EARLIER APPLICATION NUMBER: 09/328,111
; EARLIER FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: 60/117,393
; EARLIER FILING DATE: 1999-01-27
; EARLIER APPLICATION NUMBER: 60/098,639
; EARLIER FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 544
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 220
; LENGTH: 616
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(616)
; OTHER INFORMATION: n = A,T,C or G
US-09-385-982-220
```

```
Query Match          1.1%; Score 19; DB 4; Length 616;
Best Local Similarity 100.0%; Pred. No. 9.9;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 617 TGATGACGTGCTTCAC 635
Db 127 TGATGACGTGCTTCACC 109
```

RESULT 9  
US-09-149-476-225/c  
Sequence 225, Application US/09149476  
Patent No. 6420526  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: 186 Human Secreted proteins.  
FILE REFERENCE: P2002P1  
CURRENT APPLICATION NUMBER: US/09/149,476  
CURRENT FILING DATE: 1998-09-08  
EARLIER APPLICATION NUMBER: PCT/US98/04493  
EARLIER FILING DATE: 1998-03-06  
EARLIER APPLICATION NUMBER: 60/040,162  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,333  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/038,621  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,626  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,334  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,336  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,163  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/047,600  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,615  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,597  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,502  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,633  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,583  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,617  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,618  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,503  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,592  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,581  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,584  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,500  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,587  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,492  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,598  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,613  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,582  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,596  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,612  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,632  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,601  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/043,580  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,568  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,314  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,569  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,311  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,671  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,674  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,669  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,312  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,313  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,672  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,315  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/048,974  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/056,886  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,877  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,889  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,893  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,630  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,878  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,662  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,872  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,882  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,637  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,903  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,888  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,879  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,880  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,894  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,911  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,636  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,874  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,910  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,864  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,631  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,845  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,892  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/057,761  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/047,595

EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,599  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,588  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,585  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,586  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,590  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,594  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,589  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,593  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,614  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/043,578  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,576  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/047,501  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/043,670  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/056,632  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,664  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,876  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,881  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,909  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,875  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,862  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,887  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,908  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/048,964  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/057,650  
EARLIER FILING DATE: 1997-09-05  
EARLIER APPLICATION NUMBER: 60/056,884  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/057,669  
EARLIER FILING DATE: 1997-09-05  
EARLIER APPLICATION NUMBER: 60/049,610  
EARLIER FILING DATE: 1997-06-13  
EARLIER APPLICATION NUMBER: 60/061,060  
EARLIER FILING DATE: 1997-10-02

Query Match 1.1%; Score 19; DB 4; Length 1196;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 19; Conservative 0; Mismatches 0;  
Indels 0; Gaps 0;

OY 617 TGATGACGGTGGCTTACC 635  
|||||  
DB 134 TGATGACGGTGGCTTACC 116

RESULT 10  
US-09-149-476-57/c  
Sequence 57, Application US/09149476  
Patent No. 6420526  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.

TITLE OF INVENTION: 186 Human Secreted proteins  
FILE REFERENCE: P2002p1  
CURRENT APPLICATION NUMBER: US/09/149,476  
CURRENT FILING DATE: 1998-09-08  
EARLIER APPLICATION NUMBER: PCT/US98/04493  
EARLIER FILING DATE: 1998-03-06  
EARLIER APPLICATION NUMBER: 60/040,162  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,333  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/038,621  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,626  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,334  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,336  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,163  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/047,600  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,615  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,597  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,502  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,633  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,583  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,617  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,618  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,503  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,592  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,581  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,584  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,500  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,587  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,492  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,598  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,613  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,582  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,596  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,612  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,632  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,601  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/043,580  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,568  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,314  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,569  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,311

EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,671  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,674  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,669  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,312  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,313  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,672  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,315  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/048,974  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/056,886  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,877  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,889  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,893  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,630  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,878  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,662  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,872  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,882  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,637  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,903  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,888  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,879  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,880  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,894  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,911  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,636  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,874  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,910  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,864  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,631  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,845  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,892  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/057,761  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/047,595  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,599  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,588  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,585  
EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,586  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,590  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,594  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,589  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,593  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,614  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/043,578  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,576  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/047,501  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/043,670  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/056,632  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,664  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,876  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,881  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,909  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,875  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,862  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,887  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,908  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/048,964  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/057,650  
EARLIER FILING DATE: 1997-09-05  
EARLIER APPLICATION NUMBER: 60/056,884  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/057,669  
EARLIER FILING DATE: 1997-09-05  
EARLIER APPLICATION NUMBER: 60/049,610  
EARLIER FILING DATE: 1997-06-13  
EARLIER APPLICATION NUMBER: 60/061,060  
EARLIER FILING DATE: 1997-10-02

Query Match 1.1%; Score 19; DB 4; Length 1220;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 617 TGATGACGGTGGCTTCACC 635  
Db 128 TGATGACGGTGGCTTCACC 110

RESULT 11  
US-09-213-768-1/c  
Sequence 1, Application US/09213768  
Patent No. 5985664  
GENERAL INFORMATION:  
APPLICANT: Brenda F. Baker  
APPLICANT: Lex M. Cowsett  
TITLE OF INVENTION: ANTISENSE MODULATION OF SENTRIN EXPRESSION  
FILE REFERENCE: RTS-0026  
CURRENT APPLICATION NUMBER: US/09/213,768  
CURRENT FILING DATE: 1998-12-17  
NUMBER OF SEQ ID NOS: 47  
SEQ ID NO 1

LENGTH: 1514  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (136)..(441)  
US-09-213-768-1

Query Match 1.1%; Score 19; DB 2; Length 1514;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 617 TGATGACGTGTGCTTCAC 635  
|||||  
DB 136 TGATGACGTGTGCTTCAC 118

RESULT 12  
US-09-668-680-13  
Sequence 13, Application US/09668680  
Patent No. 6436703  
GENERAL INFORMATION:  
APPLICANT: Tang, Y. Tom  
APPLICANT: Liu, Chenghua  
APPLICANT: Zhou, Ping  
APPLICANT: Asundi, Vinod  
APPLICANT: Zhang, Jie  
APPLICANT: Wang, Jian-Rui  
APPLICANT: Xue, Aidong J.  
APPLICANT: Xu, Chongjun  
APPLICANT: Drmanac, Radoje T.  
TITLE OF INVENTION: No. 6436703el Nucleic Acids and  
FILE OF INVENTION: Polypeptides  
FILE REFERENCE: 790CIP2A  
CURRENT APPLICATION NUMBER: US/09/668, 680  
CURRENT FILING DATE: 2000-09-22  
PRIOR APPLICATION NUMBER: 09/649,167  
PRIOR FILING DATE: 2000-08-23  
PRIOR APPLICATION NUMBER: 09/540,217  
PRIOR FILING DATE: 2000-03-31  
NUMBER OF SEQ ID NOS: 13  
SOFTWARE: PL\_FL\_genes Version 2.0  
SEQ ID NO 13  
LENGTH: 1539  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (130)..(1539)  
US-09-668-680-13

Query Match 1.1%; Score 19; DB 4; Length 1539;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 287 CTCCTCTGCTCTCTCA 305  
|||||  
DB 774 CTCCTCTGCTCTCTCA 792

RESULT 13  
US-09-213-768-2/c  
Sequence 2, Application US/09213768  
Patent No. 5985664  
GENERAL INFORMATION:  
APPLICANT: Brenda F. Baker  
APPLICANT: Lex M. Cowsett  
TITLE OF INVENTION: ANTISENSE MODULATION OF SENTRIN EXPRESSION  
FILE REFERENCE: RUS-0026  
CURRENT APPLICATION NUMBER: US/09/213,768  
CURRENT FILING DATE: 1998-12-17  
NUMBER OF SEQ ID NOS: 47  
SEQ ID NO 2

LENGTH: 20  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: PCR Primer  
US-09-213-768-2

Query Match 1.1%; Score 18; DB 2; Length 20;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 617 TGATGACGTGTGCTTCAC 634  
|||||  
DB 18 TGATGACGTGTGCTTCAC 1

RESULT 14  
US-09-313-930-1/c  
Sequence 1, Application US/09313930  
Patent No. 6235723  
GENERAL INFORMATION:  
APPLICANT: Dean, Nicholas M.  
TITLE OF INVENTION: Antisense Oligonucleotide Modulation of Human Protein  
FILE REFERENCE: ISPH-0357  
CURRENT APPLICATION NUMBER: US/09/313,930  
CURRENT FILING DATE: 1999-05-18  
NUMBER OF SEQ ID NOS: 20  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 1  
LENGTH: 2104  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (59)..(2089)  
PUBLICATION INFORMATION:  
AUTHORS: Aris, J. P.  
AUTHORS: Basta, P. V.  
AUTHORS: Holmes, W. D.  
AUTHORS: Ballas, L. M.  
AUTHORS: Moornaw, C.  
AUTHORS: Rankl, N. B.  
AUTHORS: Blobel, G.  
AUTHORS: Loomis, C. R.  
AUTHORS: Burns, D. J.  
TITLE: Molecular and biochemical characterization of a  
JOURNAL: Biochim. Biophys. Acta  
VOLUME: 1174  
ISSUE: 2  
PAGES: 171-181  
DATE: 1993-08-19  
DATABASE ACCESSION NUMBER: L07860  
DATABASE ENTRY DATE: 1993-11-02  
US-09-313-930-1

Query Match 1.1%; Score 18; DB 4; Length 2104;  
Best Local Similarity 100.0%; Pred. No. 32;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1658 GCAGAGTACTCTCTCTG 1675  
|||||  
DB 1185 GCAGAGTACTCTCTCTG 1168

RESULT 15  
US-08-259-451-10  
Sequence 10, Application US/08259451  
Patent No. 6406841  
GENERAL INFORMATION:  
APPLICANT: Lee, Helen H.  
APPLICANT: Swanson, Priscilla A.

APPLICANT: Idler, Kenneth B.  
APPLICANT: Rosenblatt, Joseph D.  
APPLICANT: Chen, Irvin S. Y.  
APPLICANT: Golde, David W.  
APPLICANT: Robertson, Eugene F.  
APPLICANT: Stephens, John E.  
APPLICANT: Chan, Emerson W.  
APPLICANT: Buytendorp, Mark H.  
APPLICANT: Johnson, Joan E.  
APPLICANT: Moley, Cheryl T.  
APPLICANT: Peterson, Bryan  
APPLICANT: Edwards, Michelle  
APPLICANT: Guidinger, Peggy  
APPLICANT: Tale, Cynthia  
TITLE OF INVENTION: HTLV-III/RA Compositions  
TITLE OF INVENTION: and Assays for Detecting HTLV Infection  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Abbott Laboratories  
STREET: One Abbott Park Road  
CITY: Abbott Park  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60064  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb  
MEDIUM TYPE: storage  
COMPUTER: IBM  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WordPerfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/259,451  
FILING DATE: 20-JUN-1994  
CLASSIFICATION: 436  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/086,415  
FILING DATE: 01-JUL-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Daniel W. Collins  
REGISTRATION NUMBER: 31,912  
REFERENCE/DOCKET NUMBER: 5381.US.P1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (708) 937-6365  
TELEFAX: (708) 938-2623  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2949 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single stranded  
TOPOLOGY: linear  
US-08-259-451-10

Query Match 1.1%; Score 18; DB 4; length 2949;  
Best Local Similarity 100.0%; Pred. No. 32;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 GGCTTCACGGCTCACC 210  
|||||  
Db 2071 GGCTTCACGGCTCACC 2088

Search completed: July 28, 2003, 15:58:52  
Job time : 94.3356 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 14:30:21 ; Search time 420.627 Seconds  
(without alignments)  
8377.033 Million cell updates/sec

Title: US-10-010-408-1

Perfect score: 1708  
Sequence: 1 GAGCGTCTGATCTCAGAG.....GCTTAGATAAACCCCAA 1708

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 1439767 seqs, 1031500376 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2879534

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published\_Applications\_NA:\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/2/pubpna/PCR\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/2/pubpna/PCRUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
10: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq1:\*  
11: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq2:\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq3:\*  
13: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
15: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*  
16: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*  
17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID               | Description        |
|------------|-------|-------------|--------|---------------------|--------------------|
| 1          | 1708  | 100.0       | 1708   | US-10-010-408-1     | Sequence 1, Appli  |
| 2          | 753   | 44.1        | 753    | US-10-010-408-3     | Sequence 3, Appli  |
| 3          | 681   | 39.9        | 681    | US-10-010-408-12    | Sequence 12, Appli |
| 4          | 354   | 20.7        | 439    | US-09-956-622A-23   | Sequence 23, Appli |
| 5          | 210   | 12.3        | 210    | US-10-010-408-8     | Sequence 8, Appli  |
| 6          | 177   | 10.4        | 177    | US-10-010-408-5     | Sequence 5, Appli  |
| 7          | 174   | 10.2        | 174    | US-10-010-408-10    | Sequence 10, Appli |
| 8          | 90    | 5.3         | 1734   | US-10-112-267-17    | Sequence 17, Appli |
| 9          | 90    | 5.3         | 1734   | US-10-112-267-18    | Sequence 18, Appli |
| 10         | 32    | 1.9         | 199    | US-09-864-761-23432 | Sequence 23432, A  |
| 11         | 32    | 1.9         | 586    | US-09-864-761-6698  | Sequence 6698, Ap  |
| 12         | 32    | 1.9         | 738    | US-10-112-267-38    | Sequence 38, Appli |
| 13         | 32    | 1.9         | 841    | US-10-112-267-39    | Sequence 39, Appli |
| 14         | 32    | 1.9         | 1266   | US-10-137-866-319   | Sequence 319, App  |
| 15         | 32    | 1.9         | 1266   | US-10-146-726-319   | Sequence 319, App  |
| 16         | 32    | 1.9         | 1266   | US-10-146-727-319   | Sequence 319, App  |

|    |    |     |      |    |                   |                   |
|----|----|-----|------|----|-------------------|-------------------|
| 17 | 32 | 1.9 | 1266 | 14 | US-10-146-788-319 | Sequence 319, App |
| 18 | 32 | 1.9 | 1266 | 14 | US-10-152-380-319 | Sequence 319, App |
| 19 | 32 | 1.9 | 1266 | 14 | US-10-153-934-319 | Sequence 319, App |
| 20 | 32 | 1.9 | 1266 | 15 | US-10-028-072-319 | Sequence 319, App |
| 21 | 32 | 1.9 | 1266 | 15 | US-10-121-049-319 | Sequence 319, App |
| 22 | 32 | 1.9 | 1266 | 15 | US-10-123-904-319 | Sequence 319, App |
| 23 | 32 | 1.9 | 1266 | 15 | US-10-140-470-319 | Sequence 319, App |
| 24 | 32 | 1.9 | 1266 | 15 | US-10-175-746-319 | Sequence 319, App |
| 25 | 32 | 1.9 | 1266 | 15 | US-10-176-918-319 | Sequence 319, App |
| 26 | 32 | 1.9 | 1266 | 15 | US-10-176-821-319 | Sequence 319, App |
| 27 | 32 | 1.9 | 1266 | 15 | US-10-137-865-319 | Sequence 319, App |
| 28 | 32 | 1.9 | 1266 | 15 | US-10-140-474-319 | Sequence 319, App |
| 29 | 32 | 1.9 | 1266 | 15 | US-10-142-431-319 | Sequence 319, App |
| 30 | 32 | 1.9 | 1266 | 15 | US-10-143-114-319 | Sequence 319, App |
| 31 | 32 | 1.9 | 1266 | 15 | US-10-140-002-319 | Sequence 319, App |
| 32 | 32 | 1.9 | 1266 | 15 | US-10-142-419-319 | Sequence 319, App |
| 33 | 32 | 1.9 | 1266 | 15 | US-10-123-262-319 | Sequence 319, App |
| 34 | 32 | 1.9 | 1266 | 15 | US-10-142-423-319 | Sequence 319, App |
| 35 | 32 | 1.9 | 1266 | 15 | US-10-121-050-319 | Sequence 319, App |
| 36 | 32 | 1.9 | 1266 | 15 | US-10-141-755-319 | Sequence 319, App |
| 37 | 32 | 1.9 | 1266 | 15 | US-10-143-032-319 | Sequence 319, App |
| 38 | 32 | 1.9 | 1266 | 15 | US-10-123-108-319 | Sequence 319, App |
| 39 | 32 | 1.9 | 1266 | 15 | US-10-123-236-319 | Sequence 319, App |
| 40 | 32 | 1.9 | 1266 | 15 | US-10-123-261-319 | Sequence 319, App |
| 41 | 32 | 1.9 | 1266 | 15 | US-10-140-921-319 | Sequence 319, App |
| 42 | 32 | 1.9 | 1266 | 15 | US-10-140-928-319 | Sequence 319, App |
| 43 | 32 | 1.9 | 1266 | 15 | US-10-121-045-319 | Sequence 319, App |
| 44 | 32 | 1.9 | 1266 | 15 | US-10-123-292-319 | Sequence 319, App |
| 45 | 32 | 1.9 | 1266 | 15 | US-10-123-903-319 | Sequence 319, App |

## ALIGNMENTS

RESULT 1  
US-10-010-408-1  
Sequence 1, Application US/10010408  
Publication No. US20020165185A1  
GENERAL INFORMATION:  
APPLICANT: John J. Castelliott, Jr.  
TITLE OF INVENTION: No. US20020165185A1 Heparin-Induced CCN-Like Molecules  
and Uses Therefor  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESSES:  
ADDRESS: LAHIVE & COCKFIELD, LLP  
STREET: 28 State Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/010,408  
FILING DATE: 07-Dec-2001  
CLASSIFICATION: <unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/044,273  
FILING DATE: March 19, 1998  
APPLICATION NUMBER: <unknown>  
FILING DATE: <unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Amy E. Mandiagouras  
REGISTRATION NUMBER: 36,207  
REFERENCE/DOCKET NUMBER: MBI-004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)742-4214  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:

LENGTH: 1708 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 249..1001  
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
 US-10-010-408-1

Query Match 100.0%; Score 1708; DB 15; Length 1708;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1708; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGGCTTCGATCTCCAGAGACCCCTGGGTGGAGAGGGCCCTTGGCAAGGTGACGCC 60  
 Db 1 GAGGCTTCGATCTCCAGAGACCCCTGGGTGGAGAGGGCCCTTGGCAAGGTGACGCC 60  
 QY 61 GCTGGCAGTGGCTTGGAAATGAGAGCTTTATTACTGGGAACCTGAGAGCTAAGAGGCTC 120  
 Db 61 GCTGGCAGTGGCTTGGAAATGAGAGCTTTATTACTGGGAACCTGAGAGCTAAGAGGCTC 120  
 QY 121 CTGTGACCTTGTTCCTAAATCTTAACTGATGCTGTGGCTTGGGCTTCAACAGCTGTACA 180  
 Db 121 CTGTGACCTTGTTCCTAAATCTTAACTGATGCTGTGGCTTGGGCTTCAACAGCTGTACA 180  
 QY 181 CACCTTCGTGTGGCCCTCCAGAGCCCTCACCTTCAGATTGAAGCTGGCTCCCAAGAGGAC 240  
 Db 181 CACCTTCGTGTGGCCCTCCAGAGCCCTCACCTTCAGATTGAAGCTGGCTCCCAAGAGGAC 240  
 QY 241 ACGGTGACATGAGAGGAGCAGCCCACTGATCCATCTTCTGGGACACTTCTCTCTGCTTC 300  
 Db 241 ACGGTGACATGAGAGGAGCAGCCCACTGATCCATCTTCTGGGACACTTCTCTCTGCTTC 300  
 QY 301 TCTCAATGATGTGTGCGGACGCTGTGGCGGACCCCTGATGCTGTCTTGGAGACACCCCC 360  
 Db 301 TCTCAATGATGTGTGCGGACGCTGTGGCGGACCCCTGATGCTGTCTTGGAGACACCCCC 360  
 QY 361 AGTGCCACAGAGGGGATACCCCTGTGTGATGATGGCTGTGGCTGTAAAGTGTGCAC 420  
 Db 361 AGTGCCACAGAGGGGATACCCCTGTGTGATGATGGCTGTGGCTGTAAAGTGTGCAC 420  
 QY 421 GGAAGCTGGGGAGTCTCTCCAGACACCTGATGTCTGCAACCCCAAGGAGGCTGTGTT 480  
 Db 421 GGAAGCTGGGGAGTCTCTCCAGACACCTGATGTCTGCAACCCCAAGGAGGCTGTGTT 480  
 QY 481 GTGAGCTGGGGAGGCGCTGTGGGGCATGGGGCTGTGTCTTGGATGAGATGAGG 540  
 Db 481 GTGAGCTGGGGAGGCGCTGTGGGGCATGGGGCTGTGTCTTGGATGAGATGAGG 540  
 QY 541 GTAGCTGTGAGTGAATGCGCCAGGTACCTGATGAGAGACCTTTAAACCAATTGCA 600  
 Db 541 GTAGCTGTGAGTGAATGCGCCAGGTACCTGATGAGAGACCTTTAAACCAATTGCA 600  
 QY 601 GGGTCTGTGCGGCTGTGTGACGGTGGCTTCACTGCGCCGCGCTGTGACGTAGAGTG 660  
 Db 601 GGGTCTGTGCGGCTGTGTGACGGTGGCTTCACTGCGCCGCGCTGTGACGTAGAGTG 660  
 QY 661 TGGGGCTGGCGAGCTGGGACCTGCCAGCGCCCAAGGATACAGGTGGCCAGGAAGTGT 720  
 Db 661 TGGGGCTGGCGAGCTGGGACCTGCCAGCGCCCAAGGATACAGGTGGCCAGGAAGTGT 720  
 QY 721 GCCCGGAGTGGTATGTGACAGAGGATGACACCGGCGATCCAGAGGCTCCAGGCGCAAG 780  
 Db 721 GCCCGGAGTGGTATGTGACAGAGGATGACACCGGCGATCCAGAGGCTCCAGGCGCAAG 780  
 QY 781 GACACCACTTCTGCGCTTGTACTCTGCTGTGCTGATGCTCTCTGTGTCCAAATTTGA 840  
 Db 781 GACACCACTTCTGCGCTTGTACTCTGCTGTGCTGATGCTCTCTGTGTCCAAATTTGA 840  
 QY 841 GCACAGCCTGGGGCGGCTGTCAACCACTGTGGGTGGGATAGCCACCCAGTGTCA 900  
 Db 841 GCACAGCCTGGGGCGGCTGTCAACCACTGTGGGTGGGATAGCCACCCAGTGTCA 900

Db 841 GCACAGCCTGGGGCGGCTGTCAACCACTGTGGGTGGGATAGCCACCCAGTGTCA 900  
 QY 901 ACCAGAACGATTTGTGCCAATGTGAGATTCACAGCGGCCCTGTGTGCCACAGACCTGGC 960  
 Db 901 ACCAGAACGATTTGTGCCAATGTGAGATTCACAGCGGCCCTGTGTGCCACAGACCTGGC 960  
 QY 961 TGGCAGCCAGGAGCCACAGACCTCATGGAACAGTGTCTTAAAGCCAACTGGGGATGCGGA 1020  
 Db 961 TGGCAGCCAGGAGCCACAGACCTCATGGAACAGTGTCTTAAAGCCAACTGGGGATGCGGA 1020  
 QY 1021 TACAGGGCTGCCATCTCTCAGCAATGACACCTCAGACAGGCGCTGAGTGTGATAT 1080  
 Db 1021 TACAGGGCTGCCATCTCTCAGCAATGACACCTCAGACAGGCGCTGAGTGTGATAT 1080  
 QY 1081 GCTCTTCTCATGCTCTTGGCTGACATTAAGTGTCTGCTGCTGATTCAGTGTAGAGCC 1140  
 Db 1081 GCTCTTCTCATGCTCTTGGCTGACATTAAGTGTCTGCTGCTGATTCAGTGTAGAGCC 1140  
 QY 1141 ACTGAGCCATCCCTGCTGTGTGAGTGAAGCGGAGACAGGTACCAAGTCCAGTCTCTG 1200  
 Db 1141 ACTGAGCCATCCCTGCTGTGTGAGTGAAGCGGAGACAGGTACCAAGTCCAGTCTCTG 1200  
 QY 1201 GTTCAGCCGGAATCTGGGCTTCTGCTGCTGCTATTCCTCAAAACATCCCTGTACAAAAG 1260  
 Db 1201 GTTCAGCCGGAATCTGGGCTTCTGCTGCTGCTATTCCTCAAAACATCCCTGTACAAAAG 1260  
 QY 1261 GACACCAAAAAGACCTTTAAACCTAAGCTATATGAGGCAACCTGCGCACCGCTGTGG 1320  
 Db 1261 GACACCAAAAAGACCTTTAAACCTAAGCTATATGAGGCAACCTGCGCACCGCTGTGG 1320  
 QY 1321 GATTAAGTCAATGTTAAGACACAGACAGATGCTGTAATCTTCAATTCCTCTTGG 1380  
 Db 1321 GATTAAGTCAATGTTAAGACACAGACAGATGCTGTAATCTTCAATTCCTCTTGG 1380  
 QY 1381 ACTTCGTATGCTTGTCCCAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440  
 Db 1381 ACTTCGTATGCTTGTCCCAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440  
 QY 1441 TGAGAACACCCCTGCTGTGCGGGAATTTCAAGGGGAGAAATCTCTGTGAACATGAAGA 1500  
 Db 1441 TGAGAACACCCCTGCTGTGCGGGAATTTCAAGGGGAGAAATCTCTGTGAACATGAAGA 1500  
 QY 1501 GATGAATCACACTGTCTTAAAGAAATTCCTGAAGTCCAGAACTTGAGCTTTGATTTT 1560  
 Db 1501 GATGAATCACACTGTCTTAAAGAAATTCCTGAAGTCCAGAACTTGAGCTTTGATTTT 1560  
 QY 1561 CAGGAATGACATCTCTTAAGACATGCGAAACAGGAAGGCTCCACACTCTGGGAGGCC 1620  
 Db 1561 CAGGAATGACATCTCTTAAGACATGCGAAACAGGAAGGCTCCACACTCTGGGAGGCC 1620  
 QY 1621 AGGGCTTCTCTTCCAGCATGAGAAAGACAAAGGACAGCAGATCTCTCTGTGAGAGA 1680  
 Db 1621 AGGGCTTCTCTTCCAGCATGAGAAAGACAAAGGACAGCAGATCTCTCTGTGAGAGA 1680  
 QY 1681 CTAGTCTAGCCTTGAATAAACCCCAA 1708  
 Db 1681 CTAGTCTAGCCTTGAATAAACCCCAA 1708

RESULT 2  
 US-10-010-408-3  
 ; Sequence 3, Application US/10010408  
 ; Publication No. US20020165185A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: John J. Castellot, Jr.  
 ; TITLE OF INVENTION: NO. US20020165185A1el Heparin-Induced CCN-Like Molecules  
 ; and Uses Therefor  
 ; NUMBER OF SEQUENCES: 13  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: LAHIVE & COCKFIELD, LLP  
 ; STREET: 28 State Street  
 ; CITY: Boston  
 ; STATE: Massachusetts

COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/010,408  
FILING DATE: 07-Dec-2001  
CLASSIFICATION: <unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/044,273  
FILING DATE: March 19, 1998  
APPLICATION NUMBER: <unknown>  
FILING DATE: <unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Amy E. Mandragouras  
REGISTRATION NUMBER: 36,207  
REFERENCE/DOCKET NUMBER: MBI-004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)742-4214  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 753 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..750  
SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
US-10-010-408-3

Query Match 44.1%; Score 753; DB 15; Length 753;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 249 ATGAGGGGACCCACCTGATCCATCTTCTGSCCACTTCTCTGCTTCTCTCAATG 308  
DB 1 ATGAGGGGACCCACCTGATCCATCTTCTGSCCACTTCTCTGCTTCTCTCAATG 60  
QY 309 GTGATGCGCAGCTGTGCGGACACCCCTGCTGCTGAGACACCCACGAGTGGCCA 368  
DB 61 GTGATGCGCAGCTGTGCGGACACCCCTGCTGCTGAGACACCCACGAGTGGCCA 120  
QY 369 CAGGGGATACCCCTGTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 428  
DB 121 CAGGGGATACCCCTGTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180  
QY 429 GGGAGTCTCGACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 488  
DB 181 GGGAGTCTCGACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240  
QY 489 GGGGAGGCGCTGGGGGCGCATGGGGCTGTGTCTCTGAGAGAGAGAGAGAGAGCT 548  
DB 241 GGGGAGGCGCTGGGGGCGCATGGGGCTGTGTCTCTGAGAGAGAGAGAGAGAGCT 300  
QY 549 GAGTGATGCGCAGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 608  
DB 301 GAGTGATGCGCAGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 360  
QY 609 TGCCGCTGTGATGACGGTGTCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 668  
DB 361 TGCCGCTGTGATGACGGTGTCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
QY 669 CCCAGCTGGAGTGGCCAGGCGCCCAAGAGATACAGGTGCGAGGAAATGCTGCCCGGAG 728  
DB 421 CCCAGCTGGAGTGGCCAGGCGCCCAAGAGATACAGGTGCGAGGAAATGCTGCCCGGAG 480  
QY 729 TGGGTATGTGACACAGGAGTGAACACGGCGATCCAGCGCTCCACGCGCAAGAGACCAA 788

DB 481 TGGGTATGTGACACAGGAGTGAACACGGCGATCCAGCGCTCCACGCGCAAGAGACCAA 540  
QY 789 CTTTCTGCGCTTGTGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 848  
DB 541 CTTTCTGCGCTTGTGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
QY 849 TGGGGCGCGCTGCTGACACCTGTGGGCTGGGCTGAGTGAACACCGAGTGTCCACAGAAC 908  
DB 601 TGGGGCGCGCTGCTGACACCTGTGGGCTGGGCTGAGTGAACACCGAGTGTCCACAGAAC 660  
QY 909 CGATTCTGCGCAACTGAGATGCCAAGCGCGCTGTGTGTGCGCAGACCTGCTGCGAGCC 968  
DB 661 CGATTCTGCGCAACTGAGATGCCAAGCGCGCTGTGTGTGCGCAGACCTGCTGCGAGCC 720  
QY 969 AGGAGCCACAGCTCATGAGACAGTCTTTCTAA 1001  
DB 721 AGGAGCCACAGCTCATGAGACAGTCTTTCTAA 753

RESULT 3  
US-10-010-408-12  
Sequence 12, Application US/10010408  
Publication No. US20020165185A1  
GENERAL INFORMATION:  
APPLICANT: John J. Castelli et al.  
TITLE OF INVENTION: No. US20020165185A1el Heparin-Induced CCN-Like Molecules  
and Uses Thereof  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD, LLP  
STREET: 28 State Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/010,408  
FILING DATE: 07-Dec-2001  
CLASSIFICATION: <unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/044,273  
FILING DATE: March 19, 1998  
APPLICATION NUMBER: <unknown>  
FILING DATE: <unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Amy E. Mandragouras  
REGISTRATION NUMBER: 36,207  
REFERENCE/DOCKET NUMBER: MBI-004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)742-4214  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 681 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..681  
SEQUENCE DESCRIPTION: SEQ ID NO: 12:  
US-10-010-408-12  
Query Match 39.9%; Score 681; DB 15; Length 681;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 681; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 318 CACCTGCGGACACCCCTGTACCTGTCTTGGACACACCCAGTCCACAGGGGTA 377  
 Db 1 CACCTGCGGACACCCCTGTACCTGTCTTGGACACACCCAGTCCACAGGGGTA 60  
 QY 378 CCCCTGGTGTGATGGCTGTGGCTGTGTAAGTGTGTGACAGGAGGCTGGGGAGTCC 437  
 Db 61 CCCCTGGTGTGATGGCTGTGGCTGTGTAAGTGTGTGACAGGAGGCTGGGGAGTCC 120  
 QY 438 TGGGACACCTGATGTCTGACACCCAGCAGGAGGCTGGTTTGTACGCTGGGGAGGC 497  
 Db 121 TGGGACACCTGATGTCTGACACCCAGCAGGAGGCTGGTTTGTACGCTGGGGAGGC 180  
 QY 498 CCGTGGGCGCATGGGGCTGTGTCTCTTGGATGAGATGACGATGCTGTGAGTGAAT 557  
 Db 181 CCGTGGGCGCATGGGGCTGTGTCTCTTGGATGAGATGACGATGCTGTGAGTGAAT 240  
 QY 558 GGGCGCAGGTACTGTGATGAGAGACCTTTAAACCCATTTGCAAGGCTCTGTGCGCTGT 617  
 Db 241 GGGCGCAGGTACTGTGATGAGAGACCTTTAAACCCATTTGCAAGGCTCTGTGCGCTGT 300  
 QY 618 GATGACGCTGGCTTACCTGCTGCTGCGCTGTGCGAGTGAATGAGGCTGGCCAGCTGG 677  
 Db 301 GATGACGCTGGCTTACCTGCTGCTGCGCTGTGCGAGTGAATGAGGCTGGCCAGCTGG 360  
 QY 678 GACTGCCACGCGCCCAAGAGATACAGGTCCAGGAAAGCTGCGCCGAGTGGTATGT 737  
 Db 361 GACTGCCACGCGCCCAAGAGATACAGGTCCAGGAAAGCTGCGCCGAGTGGTATGT 420  
 QY 738 GACCAGGAGTGAACCGCGGATCCAGCGCTCCACGCGCGCAAGAGACCACTTTTGGCC 797  
 Db 421 GACCAGGAGTGAACCGCGGATCCAGCGCTCCACGCGCGCAAGAGACCACTTTTGGCC 480  
 QY 798 CTTGTCACTCTCTCTCTGTGTGATGCTCTTGTCAAAATGGAGACAGCTGGGGCCC 857  
 Db 481 CTTGTCACTCTCTCTCTGTGTGATGCTCTTGTCAAAATGGAGACAGCTGGGGCCC 540  
 QY 858 TGGTCAACACCTGTGGGCTGGGCAATAGCACCAGAGTGTCAACCAAGACCAATTCG 917  
 Db 541 TGGTCAACACCTGTGGGCTGGGCAATAGCACCAGAGTGTCAACCAAGACCAATTCG 600  
 QY 918 CACTGAGATCCACGCGCGCTGTGTCTCCAGACCGCTGGCAGCAGGAGGAC 977  
 Db 601 CACTGAGATCCACGCGCGCTGTGTCTCCAGACCGCTGGCAGCAGGAGGAC 660  
 QY 978 AGCTCATGAACAGTCTTC 998  
 Db 661 AGCTCATGAACAGTCTTC 681

RESULT 4  
 US-09-956-622A-23  
 ; Sequence 23, Application US/09956622A  
 ; Publication No. US20030091973A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Horesovsky, Gregory J  
 ; APPLICANT: Raha, Debashish  
 ; TITLE OF INVENTION: Method of Identifying Osteoregenerative Agents Using  
 ; TITLE OF INVENTION: Differential Gene Expression  
 ; FILE REFERENCE: 21402-445  
 ; CURRENT APPLICATION NUMBER: US/09/956,622A  
 ; PRIOR FILING DATE: 2001-09-19  
 ; PRIOR APPLICATION NUMBER: 60/233,579  
 ; NUMBER OF SEQ ID NOS: 53  
 ; SOFTWARE: Patent In Ver. 2.1  
 ; SEQ ID NO 23  
 ; LENGTH: 439  
 ; TYPE: DNA  
 ; ORGANISM: Rattus norvegicus  
 US-09-956-622A-23

Query Match 20.7%; Score 354; DB 12; Length 439;  
 Best Local Similarity 99.8%; Pred. No. 3e-178;  
 Matches 404; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1250 TGTACAAAAGAGCAACCAAAAGACCTTTAAACCTAGGCTATACGTGGCAACCTGGCC 1309  
 Db 1 TGTACAAAAGAGCAACCAAAAGACCTTTAAACCTAGGCTATACGTGGCAACCTGGCC 60  
 QY 1310 ACCGTGCTGGGAGTAAGTCAATGTTAGACACAGACAGATGTGCTGAACTTCCAT 1369  
 Db 61 ACCGTGCTGGGAGTAAGTCAATGTTAGACACAGACAGATGTGCTGAACTTCCAT 120  
 QY 1370 TCCCTCTTGGACTTCTGTATGCTTGTCCCAAGATGATGAATGAACTGTAAAGTAC 1429  
 Db 121 TCCCTCTTGGACTTCTGTATGCTTGTCCCAAGATGATGAATGAACTGTAAAGTAC 180  
 QY 1430 CTTCCCTGACCTAGAACACCCCTGCTGCTGGGAAGTATTCAGGGGAGAAATTCCTGT 1489  
 Db 181 CTTCCCTGACCTAGAACACCCCTGCTGCTGGGAAGTATTCAGGGGAGAAATTCCTGT 240  
 QY 1490 GAACATGAAGATGAATCACTGTCTTGAAGAAATTCCTGAAGTCCAGAACTTGAG 1549  
 Db 241 GAACATGAAGATGAATCACTGTCTTGAAGAAATTCCTGAAGTCCAGAACTTGAG 300  
 QY 1550 CTTTGTATTTTCAGAAATGCATCTCTTAAGCACTGCGAAACAGGAAGCTTCACACC 1609  
 Db 301 CTTTGTATTTTCAGAAATGCATCTCTTAAGCACTGCGAAACAGGAAGCTTCACACC 360  
 QY 1610 TCTGGCAGCGCAGGCTTCTCTTCAAGATGAGAAAGCAAGG 1654  
 Db 361 TCTGGCAGCGCAGGCTTCTCTTCAAGATGAGAAAGCAAGG 405

RESULT 5  
 US-10-010-408-8  
 ; Sequence 8, Application US/10010408  
 ; Publication No. US2002016185A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: John J. Castelle, Jr.  
 ; TITLE OF INVENTION: No. US2002016185A1el Heparin-Induced CN-Like Molecules  
 ; and Uses Therefor  
 ; NUMBER OF SEQUENCES: 13  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: LAHIVE & COCKFIELD, LLP  
 ; STREET: 28 State Street  
 ; CITY: Boston  
 ; STATE: Massachusetts  
 ; COUNTRY: USA  
 ; ZIP: 02109  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/10/010,408  
 ; FILING DATE: 07-Dec-2001  
 ; CLASSIFICATION: <Unknown>  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 09/044,273  
 ; FILING DATE: March 19, 1998  
 ; APPLICATION NUMBER: <Unknown>  
 ; FILING DATE: <Unknown>  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Amy E. Mandragoras  
 ; REGISTRATION NUMBER: 36,207  
 ; REFERENCE/DOCKET NUMBER: MB1-004  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (617)227-7400  
 ; TELEFAX: (617)742-4214  
 ; INFORMATION FOR SEQ ID NO: 8:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 210 base pairs

TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..210  
SEQUENCE DESCRIPTION: SEQ ID NO: 8:  
US-10-010-408-8

Query Match 12.3%; Score 210; DB 15; Length 210;  
Best Local Similarity 100.0%; Pred. No. 2.2e-101;  
Matches 210; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 318 CAGGTGCGGACACCTGACTGCTGAGACACCCAGTCCCGCCAGAGGGGTA 377  
DB 1 CAGGTGCGGACACCTGACTGCTGAGACACCCAGTCCCGCCAGAGGGGTA 60  
QY 378 CCCCTGCTGTGATGCTGTGCTGCTGTAAGTGTGTGACAGAGGCTGGGGAGTCC 437  
DB 61 CCCCTGCTGTGATGCTGTGCTGCTGTAAGTGTGTGACAGAGGCTGGGGAGTCC 120  
QY 438 TGCACCACTGATGCTGTGCGACCCAGCCAGGCGCTGTTGTCAAGCTGGGGCAAGC 497  
DB 121 TGCACCACTGATGCTGTGCGACCCAGCCAGGCGCTGTTGTCAAGCTGGGGCAAGC 180  
QY 498 CCTGGCGGCATGGGGCTGTGTCTCTG 527  
DB 181 CTTGGCGGCATGGGGCTGTGTCTCTG 210

RESULT 6  
US-10-010-408-5  
Sequence 5, Application US/10010408  
Publication No. US20020165185A1  
GENERAL INFORMATION:  
APPLICANT: John J. Castellot, Jr.  
TITLE OF INVENTION: No. US20020165185A1el Heparin-Induced CCN-Like Molecules  
and Uses Therefor  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD, LLP  
STREET: 28 State Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/010.408  
FILING DATE: 07-Dec-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/044,273  
FILING DATE: March 19, 1998  
APPLICATION NUMBER: <Unknown>  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Amy E. Mandragouras  
REGISTRATION NUMBER: 36,207  
REFERENCE/DOCKET NUMBER: MBI-004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)742-4214  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 177 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single

TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..177  
SEQUENCE DESCRIPTION: SEQ ID NO: 5:  
US-10-010-408-5

Query Match 10.4%; Score 177; DB 15; Length 177;  
Best Local Similarity 100.0%; Pred. No. 8.9e-84;  
Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 546 TGTAGGTGATGACCCAGATGACCTGATGAGAGACCTTAACCAATTGACGGTC 605  
DB 1 TGTAGGTGATGACCCAGATGACCTGATGAGAGACCTTAACCAATTGACGGTC 60  
QY 606 CTGTGCGCTGTGATGACGCTGCTTACCTGCTGCGCTGTGCACTGAGATGTCGG 665  
DB 61 CTGTGCGCTGTGATGACGCTGCTTACCTGCTGCGCTGTGCACTGAGATGTCGG 120  
QY 666 CTGCCCAGCTGGAGCTGCCACAGCCCAAGAGATTAAGTGTCCAGAAAGTCTGC 722  
DB 121 CTGCCAGCTGGAGCTGCCACAGCCCAAGAGATTAAGTGTCCAGAAAGTCTGC 177

RESULT 7  
US-10-010-408-10  
Sequence 10, Application US/10010408  
Publication No. US20020165185A1  
GENERAL INFORMATION:  
APPLICANT: John J. Castellot, Jr.  
TITLE OF INVENTION: No. US20020165185A1el Heparin-Induced CCN-Like Molecules  
and Uses Therefor  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD, LLP  
STREET: 28 State Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/010.408  
FILING DATE: 07-Dec-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/044,273  
FILING DATE: March 19, 1998  
APPLICATION NUMBER: <Unknown>  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Amy E. Mandragouras  
REGISTRATION NUMBER: 36,207  
REFERENCE/DOCKET NUMBER: MBI-004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)742-4214  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 174 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..174  
SEQUENCE DESCRIPTION: SEQ ID NO: 10:

US-10-010-408-10

Query Match 10.2% Score 174; DB 15; Length 174;  
Best Local Similarity 100.0%; Pred. No. 3.5e-82;  
Matches 174; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 825 CCTGTGCAAAATGGAGACAGACCTGGGGCCCCCTGCTCAACCACTGTGGGCTGGCATA 884  
|||||  
DB 1 CCTGTGCAAAATGGAGACAGACCTGGGGCCCCCTGCTCAACCACTGTGGGCTGGCATA 60  
QY 885 GCCACCCGAGTGTCCACACGAGCGATTCCTGCAACTGAGATCCACCGCCCTGTGT 944  
|||||  
DB 61 GCCACCCGAGTGTCCACACGAGCGATTCCTGCAACTGAGATCCACCGCCCTGTGT 120  
QY 945 CTGCCAGACCTGCTGGCGAGCAGACCGACGCTCACTGGAAGAGTCTTTC 998  
|||||  
DB 121 CTGCCAGACCTGCTGGCGAGCAGACCGACGCTCACTGGAAGAGTCTTTC 174

RESULT 8

US-10-112-267-17  
Sequence 17, Application US/10112267  
Publication No. US20030068678A1

GENERAL INFORMATION:  
APPLICANT: Bolstein, David A.  
APPLICANT: Cohen, Robert  
APPLICANT: Goddard, Audrey  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: Lawrence, David A.  
APPLICANT: Levine, Arnold J.  
APPLICANT: Pennica, Diane  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME  
FILE REFERENCE: P11762  
CURRENT APPLICATION NUMBER: US/10/112,267  
PRIOR FILING DATE: 2002-03-27  
PRIOR APPLICATION NUMBER: US/09/182,145B  
PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-29  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/063,704  
PRIOR FILING DATE: EARLIER FILING DATE: 1997-10-23  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/073,612  
PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-04  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/081,695  
PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-14  
NUMBER OF SEQ ID NOS: 156  
SEQ ID NO 17  
LENGTH: 1734  
TYPE: DNA  
ORGANISM: Mus musculus  
US-10-112-267-17

Query Match 5.3% Score 90; DB 15; Length 1734;  
Best Local Similarity 100.0%; Pred. No. 2.1e-37;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 410 AGTGTGTGACGAGAGCTGGGGAGTCTGCGACCACTGCATGCTGTGGACCCAGCA 469  
|||||  
DB 418 AGTGTGTGACGAGAGCTGGGGAGTCTGCGACCACTGCATGCTGTGGACCCAGCA 477  
QY 470 GGGCCTGTTGTTCAGACCTGGGGCAGGCC 499  
|||||  
DB 478 GGGCCTGTTGTTCAGACCTGGGGCAGGCC 507

RESULT 9

US-10-112-267-18/c  
Sequence 18, Application US/10112267  
Publication No. US20030068678A1  
GENERAL INFORMATION:  
APPLICANT: Bolstein, David A.  
APPLICANT: Cohen, Robert

APPLICANT: Goddard, Audrey  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: Lawrence, David A.  
APPLICANT: Levine, Arnold J.  
APPLICANT: Pennica, Diane  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME  
FILE REFERENCE: P11762  
CURRENT APPLICATION NUMBER: US/10/112,267  
PRIOR FILING DATE: 2002-03-27  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/182,145B  
PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-29  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/063,704  
PRIOR FILING DATE: EARLIER FILING DATE: 1997-10-23  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/073,612  
PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-04  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/081,695  
PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-14  
NUMBER OF SEQ ID NOS: 156  
SEQ ID NO 18  
LENGTH: 1734  
TYPE: DNA  
ORGANISM: Mus musculus  
US-10-112-267-18

Query Match 5.3% Score 90; DB 15; Length 1734;  
Best Local Similarity 100.0%; Pred. No. 2.1e-37;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 410 AGTGTGTGACGAGAGCTGGGGAGTCTGCGACCACTGCATGCTGTGGACCCAGCA 469  
|||||  
DB 1317 AGTGTGTGACGAGAGCTGGGGAGTCTGCGACCACTGCATGCTGTGGACCCAGCA 1258  
QY 470 GGGCCTGTTGTTCAGACCTGGGGCAGGCC 499  
|||||  
DB 1257 GGGCCTGTTGTTCAGACCTGGGGCAGGCC 1228

RESULT 10

US-09-864-761-23432  
Sequence 23432, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: Aeonica-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
PRIOR FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665

```

: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00668
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00663
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00662
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00661
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00670
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: US 60/234,687
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: US 09/608,408
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/774,203
: PRIOR FILING DATE: 2001-01-29
: NUMBER OF SEQ ID NOS: 49117
: SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
: SEQ ID NO 23432
: LENGTH: 199
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: MAP TO AL139352.8
: OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2
: OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.7
: OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.9
: OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.8
: OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.8
: OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.7
: OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.7
: OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.7
: OTHER INFORMATION: NT HIT: AF083500.1, EVALU 1.00e-108
: OTHER INFORMATION: SWISSPROT HIT: O19113, EVALU 9.00e-19
US-09-864-761-23432
Query Match 1.9%; Score 32; DB 10; Length 199;
Best Local Similarity 100.0%; Pred. No. 2,1e-06;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 654 GAGGATGTGCGGCTGCCAGCTGGGACTGCC 685
Db 129 GAGGATGTGCGGCTGCCAGCTGGGACTGCC 160

RESULT 11
US-09-864-761-6698
: Sequence 6698, Application US/09864761
: Patent No. US20020048763A1
: GENERAL INFORMATION:
: APPLICANT: Penn, Sharon G.
: APPLICANT: Rank, David R.
: APPLICANT: Hanzel, David K.
: APPLICANT: Chen, Wensheng
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
: FILE REFERENCE: Aegm1ca-X-1
: CURRENT APPLICATION NUMBER: US/09/864,761
: PRIOR FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 60/180,312
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/632,366
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: GB 24263,6
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 60/236,359
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/00666
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00667
```

```

: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00664
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00669
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00665
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00668
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00663
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00662
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00661
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00670
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: US 60/234,687
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: US 09/608,408
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/774,203
: PRIOR FILING DATE: 2001-01-29
: NUMBER OF SEQ ID NOS: 49117
: SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
: SEQ ID NO 6698
: LENGTH: 586
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: MAP TO AL139352.8
: OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2
: OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.7
: OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.9
: OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.8
: OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.8
: OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.7
: OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.7
: OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.7
US-09-864-761-6698
Query Match 1.9%; Score 32; DB 10; Length 586;
Best Local Similarity 100.0%; Pred. No. 2e-06;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 654 GAGGATGTGCGGCTGCCAGCTGGGACTGCC 685
Db 342 GAGGATGTGCGGCTGCCAGCTGGGACTGCC 373

RESULT 12
US-10-112-267-38
: Sequence 38, Application US/10112267
: Publication No. US20030068678A1
: GENERAL INFORMATION:
: APPLICANT: Botstein, David A.
: APPLICANT: Cohen, Robert
: APPLICANT: Goddard, Audrey
: APPLICANT: Gurney, Austin L.
: APPLICANT: Hillan, Kenneth J.
: APPLICANT: Lawrence, David A.
: APPLICANT: Levine, Arnold J.
: APPLICANT: Penica, Diane
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Wood, William I.
: TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
: FILE REFERENCE: P1176R2
: CURRENT APPLICATION NUMBER: US/10/112,267
: PRIOR FILING DATE: 2002-03-27
: PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/182,145B
: PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-29
: PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/063,704
: PRIOR FILING DATE: EARLIER FILING DATE: 1997-10-29
```

|    |   |                     |
|----|---|---------------------|
| 1  | APPLICANT:  | Beresini, Maureen   |
| 2  | APPLICANT:  | Deforge, Laura      |
| 3  | APPLICANT:  | Desnoyers, Luc      |
| 4  | APPLICANT:  | Filvaroff, Ellen    |
| 5  | APPLICANT:  | Gao, Wei-Qiang      |
| 6  | APPLICANT:  | Gerritsen, Mary E.  |
| 7  | APPLICANT:  | Goddard, Audrey     |
| 8  | APPLICANT:  | Godowski, Paul J.   |
| 9  | APPLICANT:  | Gurney, Austin L.   |
| 10 | APPLICANT:  | Sherwood, Steven    |
| 11 | APPLICANT:  | Smith, Victoria A.  |
| 12 | APPLICANT:  | Stewart, Timothy A. |
| 13 | APPLICANT:  | Tumas, Daniel       |
| 14 | APPLICANT:  | Matanabe, Collin K  |
| 15 | APPLICANT:  | Wood, William       |
| 16 | APPLICANT:  | Zhang, Zemin        |
| 17 | TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC |                     |
| 18 | FILE OF INVENTION: ACIDS ENCODING THE SAME                              |                     |
| 19 | FILE REFERENCE: P3330R1C151   |                     |
| 20 | CURRENT APPLICATION NUMBER: US/10/137,866                               |                     |
| 21 | CURRENT FILING DATE: 2002-05-03   |                     |
| 22 | PRIOR APPLICATION NUMBER: 60/049911                                     |                     |
| 23 | PRIOR FILING DATE: 1997-06-18   |                     |
| 24 | PRIOR APPLICATION NUMBER: 60/056974                                     |                     |
| 25 | PRIOR FILING DATE: 1997-08-26   |                     |
| 26 | PRIOR APPLICATION NUMBER: 60/059113                                     |                     |
| 27 | PRIOR FILING DATE: 1997-09-17   |                     |
| 28 | PRIOR APPLICATION NUMBER: 60/059115                                     |                     |
| 29 | PRIOR FILING DATE: 1997-09-17   |                     |
| 30 | PRIOR APPLICATION NUMBER: 60/059117                                     |                     |
| 31 | PRIOR FILING DATE: 1997-09-17   |                     |
| 32 | PRIOR APPLICATION NUMBER: 60/059122                                     |                     |
| 33 | PRIOR FILING DATE: 1997-09-17   |                     |
| 34 | PRIOR APPLICATION NUMBER: 60/059184                                     |                     |
| 35 | PRIOR FILING DATE: 1997-09-17   |                     |
| 36 | PRIOR APPLICATION NUMBER: 60/059263                                     |                     |
| 37 | PRIOR FILING DATE: 1997-09-18   |                     |
| 38 | PRIOR APPLICATION NUMBER: 60/059352                                     |                     |
| 39 | PRIOR FILING DATE: 1997-09-19   |                     |
| 40 | PRIOR APPLICATION NUMBER: 60/059588                                     |                     |
| 41 | PRIOR FILING DATE: 1997-09-19   |                     |
| 42 | PRIOR APPLICATION NUMBER: 60/059836                                     |                     |
| 43 | PRIOR FILING DATE: 1997-09-24   |                     |
| 44 | PRIOR APPLICATION NUMBER: 60/062250                                     |                     |
| 45 | PRIOR FILING DATE: 1997-10-17   |                     |
| 46 | PRIOR APPLICATION NUMBER: 60/062285                                     |                     |
| 47 | PRIOR FILING DATE: 1997-10-17   |                     |
| 48 | PRIOR APPLICATION NUMBER: 60/062287                                     |                     |
| 49 | PRIOR FILING DATE: 1997-10-17   |                     |
| 50 | PRIOR APPLICATION NUMBER: 60/062814                                     |                     |
| 51 | PRIOR FILING DATE: 1997-10-24   |                     |
| 52 | PRIOR APPLICATION NUMBER: 60/062816                                     |                     |
| 53 | PRIOR FILING DATE: 1997-10-24   |                     |
| 54 | PRIOR APPLICATION NUMBER: 60/063045                                     |                     |
| 55 | PRIOR FILING DATE: 1997-10-24   |                     |
| 56 | PRIOR APPLICATION NUMBER: 60/063082                                     |                     |
| 57 | PRIOR FILING DATE: 1997-10-31   |                     |
| 58 | PRIOR APPLICATION NUMBER: 60/063127                                     |                     |
| 59 | PRIOR FILING DATE: 1997-10-24   |                     |
| 60 | PRIOR APPLICATION NUMBER: 60/063327                                     |                     |
| 61 | PRIOR FILING DATE: 1997-10-27   |                     |
| 62 | PRIOR APPLICATION NUMBER: 60/063329                                     |                     |
| 63 | PRIOR FILING DATE: 1997-10-27   |                     |
| 64 | PRIOR APPLICATION NUMBER: 60/063550                                     |                     |
| 65 | PRIOR FILING DATE: 1997-10-28   |                     |
| 66 | PRIOR APPLICATION NUMBER: 60/063561                                     |                     |
| 67 | PRIOR FILING DATE: 1997-10-28   |                     |
| 68 | PRIOR APPLICATION NUMBER: 60/063704                                     |                     |
| 69 | PRIOR FILING DATE: 1997-10-29   |                     |
| 70 | PRIOR APPLICATION NUMBER: 60/063733                                     |                     |
| 71 | PRIOR FILING DATE: 1997-10-29   |                     |
| 72 | PRIOR APPLICATION NUMBER: 60/063735                                     |                     |
| 73 | PRIOR FILING DATE: 1997-10-29   |                     |



;; PRIOR APPLICATION NUMBER: 60/063738  
;; PRIOR FILING DATE: 1997-10-29  
;; PRIOR APPLICATION NUMBER: 60/063755  
;; PRIOR FILING DATE: 1997-10-17  
;; PRIOR APPLICATION NUMBER: 60/064248  
;; PRIOR FILING DATE: 1997-11-03  
;; PRIOR APPLICATION NUMBER: 60/064809  
;; PRIOR FILING DATE: 1997-11-07  
;; PRIOR APPLICATION NUMBER: 60/065186  
;; PRIOR FILING DATE: 1997-11-12  
;; PRIOR APPLICATION NUMBER: 60/065846  
;; PRIOR FILING DATE: 1997-11-17  
;; PRIOR APPLICATION NUMBER: 60/066364  
;; PRIOR FILING DATE: 1997-11-21  
;; PRIOR APPLICATION NUMBER: 60/066453  
;; PRIOR FILING DATE: 1997-11-24  
;; PRIOR APPLICATION NUMBER: 60/066511  
;; PRIOR FILING DATE: 1997-11-24  
;; PRIOR APPLICATION NUMBER: 60/066770  
;; PRIOR FILING DATE: 1997-11-24  
;; PRIOR APPLICATION NUMBER: 60/069212  
;; PRIOR FILING DATE: 1997-12-11  
;; PRIOR APPLICATION NUMBER: 60/069278  
;; PRIOR FILING DATE: 1997-12-11  
;; PRIOR APPLICATION NUMBER: 60/069334  
;; PRIOR FILING DATE: 1997-12-11  
;; PRIOR APPLICATION NUMBER: 60/069694  
;; PRIOR FILING DATE: 1997-12-16  
;; PRIOR APPLICATION NUMBER: 60/072320  
;; PRIOR FILING DATE: 1998-01-23  
;; PRIOR APPLICATION NUMBER: 60/073612  
;; PRIOR FILING DATE: 1998-02-04  
;; PRIOR APPLICATION NUMBER: 60/074086  
;; PRIOR FILING DATE: 1998-02-09  
;; PRIOR APPLICATION NUMBER: 60/074092  
;; PRIOR FILING DATE: 1998-02-09  
;; PRIOR APPLICATION NUMBER: 60/077791  
;; PRIOR FILING DATE: 1998-03-12  
;; PRIOR APPLICATION NUMBER: 60/078910  
;; PRIOR FILING DATE: 1998-03-20  
;; PRIOR APPLICATION NUMBER: 60/079294  
;; PRIOR FILING DATE: 1998-03-25  
;; PRIOR APPLICATION NUMBER: 60/079663  
;; PRIOR FILING DATE: 1998-02-27  
;; PRIOR APPLICATION NUMBER: 60/079728  
;; PRIOR FILING DATE: 1998-03-27  
;; PRIOR APPLICATION NUMBER: 60/080165  
;; PRIOR FILING DATE: 1998-03-31  
;; PRIOR APPLICATION NUMBER: 60/081203  
;; PRIOR FILING DATE: 1998-04-09  
;; PRIOR APPLICATION NUMBER: 60/081229  
;; PRIOR FILING DATE: 1998-04-09  
;; PRIOR APPLICATION NUMBER: 60/081695  
;; PRIOR FILING DATE: 1998-04-14  
;; PRIOR APPLICATION NUMBER: 60/081817  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/081818  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/082999  
;; PRIOR FILING DATE: 1998-04-24  
;; PRIOR APPLICATION NUMBER: 60/083322  
;; PRIOR FILING DATE: 1998-04-28  
;; PRIOR APPLICATION NUMBER: 60/083545  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/084600  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084627  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084637  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/085149  
;; PRIOR FILING DATE: 1998-05-12  
;; PRIOR APPLICATION NUMBER: 60/085323

;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085338  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085339  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085579  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085697  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085704  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/086414  
;; PRIOR FILING DATE: 1998-05-22  
;; PRIOR APPLICATION NUMBER: 60/086430  
;; PRIOR FILING DATE: 1998-05-22  
;; PRIOR APPLICATION NUMBER: 60/087106  
;; PRIOR FILING DATE: 1998-05-28  
;; PRIOR APPLICATION NUMBER: 60/088026  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088730  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088741  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088810  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088858  
;; PRIOR FILING DATE: 1998-06-11  
;; PRIOR APPLICATION NUMBER: 60/089532  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089599  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089907  
;; PRIOR FILING DATE: 1998-06-18  
;; PRIOR APPLICATION NUMBER: 60/089947  
;; PRIOR FILING DATE: 1998-06-19  
;; PRIOR APPLICATION NUMBER: 60/090349  
;; PRIOR FILING DATE: 1998-06-23  
;; PRIOR APPLICATION NUMBER: 60/090429  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090445  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090538  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090863  
;; PRIOR FILING DATE: 1998-06-26  
;; PRIOR APPLICATION NUMBER: 60/091360  
;; PRIOR FILING DATE: 1998-07-01  
;; PRIOR APPLICATION NUMBER: 60/091519  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091982

Query Match 1.9%; Score 32; DB 14; Length 1266;  
Best Local Similarity 100.0%; Pred. No. 1.9e-06;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 375 GTACCCCTGCTGCTGATGCTGCTGCTGCTG 406  
|||||  
Db 136 GTACCCCTGCTGCTGATGCTGCTGCTGCTG 167

RESULT 15  
US-10-146-726-319  
;; Sequence 319, Application US/10146726  
;; Publication No. US20030129690A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Baker, Kevin P.  
;; APPLICANT: Beresini, Maureen  
;; APPLICANT: Deforge, Laura  
;; APPLICANT: Desnoyers, Luc  
;; APPLICANT: Filvaroff, Ellen  
;; APPLICANT: Gao, Wei-Qiang  
;; APPLICANT: Gerritsen, Mary E.  
;; APPLICANT: Goddard, Audrey

```

: APPLICANT: Godowski, Paul J.
: APPLICANT: Gurney, Austin L.
: APPLICANT: Sherwood, Steven
: APPLICANT: Smith, Victoria
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Watanabe, Colin K
: APPLICANT: Wood, William
: APPLICANT: Zhang, Zemin
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: FILE REFERENCE: P3330R1C308
: CURRENT APPLICATION NUMBER: US/10/146,726
: CURRENT FILING DATE: 2002-05-15
: Prior Application removed - See File Wrapper or Pajm
: NUMBER OF SEQ ID NOS: 550
: SEQ ID NO 319
: LENGTH: 1266
: TYPE: DNA
: ORGANISM: Homo Sapien
: US-10-146-726-319

```

```

Query Match      1.9%: Score 32; DB 14: length 1266;
Best Local Similarity 100.0%; Pred. No. 1.9e-06;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      375 GTACCCCTGCTGCTGATGCGTGTGCTGCTG 406
      |||
      136 GTACCCCTGCTGCTGATGCGTGTGCTGCTG 167
Db

```

Search completed: July 28, 2003, 15:36:31  
 Job time: 426.627 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 15:16:26 ; Search time 3029.91 Seconds  
(without alignments)  
9129.604 Million cell updates/sec

Title: US-10-010-408-1

Perfect score: 1708

Sequence: 1 GACGCTCTGATCTCCACAG.....GCTTAGATTAACACCCAAA 1708

Scoring table:

OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size : 0

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_trod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description        |
|------------|-------|-------------|--------|-------|--------------------|
| 1          | 501   | 29.3        | 762    | 14    | B0195526 UI-R-CNI- |
| 2          | 90    | 5.3         | 940    | 14    | B0937887           |
| 3          | 67    | 3.9         | 537    | 14    | B0560868           |
| 4          | 49    | 2.9         | 424    | 10    | BB849097           |
| 5          | 35    | 2.0         | 315    | 10    | BB374499           |
| 6          | 35    | 2.0         | 369    | 10    | BB792544           |

|    |    |     |      |    |           |
|----|----|-----|------|----|-----------|
| 7  | 35 | 2.0 | 418  | 9  | A1225477  |
| 8  | 35 | 2.0 | 488  | 9  | AA754979  |
| 9  | 35 | 2.0 | 792  | 12 | BF138093  |
| 10 | 35 | 2.0 | 966  | 12 | BF141695  |
| 11 | 33 | 1.9 | 426  | 9  | AA717584  |
| 12 | 32 | 1.9 | 380  | 12 | BG900020  |
| 13 | 32 | 1.9 | 405  | 12 | BG900069  |
| 14 | 32 | 1.9 | 489  | 14 | BM751866  |
| 15 | 32 | 1.9 | 618  | 12 | BG538695  |
| 16 | 32 | 1.9 | 620  | 13 | BG928868  |
| 17 | 32 | 1.9 | 651  | 13 | BI457141  |
| 18 | 32 | 1.9 | 749  | 9  | AL555144  |
| 19 | 32 | 1.9 | 750  | 13 | BM043988  |
| 20 | 32 | 1.9 | 790  | 13 | BM046275  |
| 21 | 32 | 1.9 | 800  | 13 | BI826781  |
| 22 | 32 | 1.9 | 886  | 13 | BI822142  |
| 23 | 32 | 1.9 | 888  | 13 | BI825652  |
| 24 | 32 | 1.9 | 916  | 13 | BI457367  |
| 25 | 32 | 1.9 | 928  | 13 | BI161474  |
| 26 | 32 | 1.9 | 933  | 14 | BQ278961  |
| 27 | 32 | 1.9 | 979  | 14 | BQ279131  |
| 28 | 32 | 1.9 | 1006 | 14 | BM921531  |
| 29 | 32 | 1.9 | 1022 | 14 | BQ952960  |
| 30 | 32 | 1.9 | 1058 | 14 | BM805088  |
| 31 | 32 | 1.9 | 1073 | 14 | BQ073722  |
| 32 | 32 | 1.9 | 1166 | 13 | BM543799  |
| 33 | 32 | 1.9 | 1251 | 14 | BQ961357  |
| 34 | 30 | 1.8 | 190  | 9  | AA647775  |
| 35 | 23 | 1.3 | 436  | 17 | AQ095651  |
| 36 | 23 | 1.3 | 742  | 13 | BI758148  |
| 37 | 23 | 1.3 | 792  | 13 | BI823598  |
| 38 | 22 | 1.3 | 307  | 10 | BB220676  |
| 39 | 21 | 1.2 | 495  | 13 | BI2004749 |
| 40 | 21 | 1.2 | 529  | 9  | AI897896  |
| 41 | 21 | 1.2 | 561  | 12 | BF051668  |
| 42 | 21 | 1.2 | 563  | 9  | AI897344  |
| 43 | 21 | 1.2 | 594  | 13 | BI682551  |
| 44 | 21 | 1.2 | 609  | 9  | AI485142  |
| 45 | 21 | 1.2 | 620  | 10 | AW223381  |

#### ALIGNMENTS

RESULT 1  
BQ195526/c 762 bp mRNA linear EST 30-APR-2002  
LOCUS  
DEFINITION  
UI-R-CNI-cmq-k-07-0-UI-s1 UI-R-CNI Rattus norvegicus cDNA clone  
BQ195526  
UI-R-CNI-cmq-k-07-0-UI 3', mRNA sequence.  
ACCESSION  
BQ195526.1 GI:20371077  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Rattus norvegicus  
Norway rat.  
Rattus norvegicus  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE  
1 (bases 1 to 762)  
Bonaldo,M.F., Lennon,G. and Soares,M.B.  
Normalization and subtraction: two approaches to facilitate gene  
discovery  
JOURNAL  
MEDLINE  
COMMENT  
Genome Res. 6 (9), 791-806 (1996)  
97044477  
Contact: Soares, MB  
Program for Rat Gene Discovery and Mapping  
University of Iowa  
451 Eckstein Medical Research Building Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: msoares@iuiiueeeg.uiowa.edu

The sequence contained an oligo-dT track that was present in the  
oligonucleotide that was used to prime the synthesis of first  
strand cDNA and therefore this may represent a bonafide poly A



```

Db      223  TGTCTTAAGAAATTCCTCAAGTCAGAACTTGAGCTTGTATTTTCAGAAATGCACA 164
Oy      1573 TCTCTTANAGCACTCGCAAAACAGAGAGCTCCACACTCTTGCGAGCGGCGCTTCTC 1632
Db      163  TCTCTTAGACACTCGCAAAACAGAGAGCTCCACACTCTTGCGAGCGGCGCTTCTC 104
Oy      1633 TTCAGCATGAGAAAGACAGAGGACAGAGTACTCTCTCTGAGGACTAGTC 1686
Db      103  TTTCAGCATGAGAAAGACAGAGGACAGAGTACTCTCTCTGAGGACTAGTC 50

RESULT 2
LOCUS   B0937887
DEFINITION AGENCOURT 8951807 NCI_CGAP_Co24 Mus musculus cDNA clone
ACCESSION B0937887
VERSION   B0937887
KEYWORDS  IMAGE:6476852 5', mRNA sequence.
SOURCE    house mouse.
ORGANISM  Mus musculus
REFERENCE 1 (bases 1 to 940)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL    Unpublished (1999)
COMMENT    Contact: Robert Strausberg, Ph.D.
            Email: cgabs-rl@mail.nih.gov
            Tissue Procurement: The Cepko Laboratory
            cDNA Library Preparation: Life Technologies, Inc.
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLNL4017 row: n column: 21
            High quality sequence stop: 543.
            Location/Qualifiers
                1..940
                /organism="Mus musculus"
                /strain="FVB/N"
                /db_xref="taxon:10090"
                /clone="IMAGE:6476852"
                /clone_lib="NCI_CGAP_Co24"
                /lab_host="DH10B (T1 phage-resistant)"
                /note="Organ: colon; Vector: PCMV-SPORT6; site_1: NotI;
                site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                Average insert size 1.6 kb. Constructed by Life
                Technologies. Note: this is a NCI CGAP Library."

BASE COUNT      169 a      277 c      288 g      200 t
ORIGIN
Query Match      5.3%; Score 90; DB 14; Length 940;
Best Local Similarity 100.0%; Pred. No. 1.8e-33;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      410 AGTGTGACACGAGAGGCTGGGGAGTCTCGACACCTGATCTGCGACCCAGCCA 469
Db      444 AGTGTGACACGAGAGGCTGGGGAGTCTCGACACCTGATCTGCGACCCAGCCA 503
Oy      470 GGGCCTGTTGTGACGCTGGGGAGGCCC 499
Db      504 GGGCCTGTTGTGACGCTGGGGAGGCCC 533

RESULT 3
LOCUS   B0560868
DEFINITION H4067A01-5 NIA Mouse 7.4K cDNA Clone Set Mus musculus cDNA clone
ACCESSION B0560868
VERSION   B0560868
KEYWORDS  537 bp mRNA linear EST 20-JUN-2002
SOURCE    H4067A01 5', mRNA sequence.
ORGANISM  Mus musculus
REFERENCE 1 (bases 1 to 537)
AUTHORS   Vanburen, V., Plao, Y., Dudekula, D.B., Qian, Y., Carter, M.G., Martin
            P.R., Stagg, C.A., Bassey, U., Alba, K., Hanatani, T., Katgul, G.J.,
            Luo, A.G. and Ko, M.S.H.
            Assembly, verification, and initial annotation of NIA 7.4K mouse
            cDNA clone set
            Unpublished (2002)
            Contact: Yong Qian
            Laboratory of Genetics
            National Institute on Aging/National Institutes of Health
            333 Cassell Drive, Suite 3000, Baltimore, MD 21224-6820, USA
            Email: cdna@igsun.grc.nia.nih.gov
            This clone set has been freely distributed to the community. Please
            visit http://lgsun.grc.nia.nih.gov/cDNA/NIA_7.4K.html for details.
            Plate: H4067 row: A column: 01
            Seq primer: -21M13 Reverse
            High quality sequence stop: 537
            POLYA=No.
            Location/Qualifiers
                1..537
                /organism="Mus musculus"
                /strain="C57BL/6"
                /db_xref="taeST:H4067A01-5"
                /db_xref="taxon:10090"
                /clone="H4067A01"
                /clone_lib="NIA Mouse 7.4K cDNA Clone Set"
                /sex="mixed"
                /dev_stage="mixed"
                /lab_host="DH10B"
                /note="Vector: pSPORT1; site_1: SalI; site_2: NotI; This
                clone is among a reartayed set of 7,407 clones from more
                than 20 cDNA libraries."

BASE COUNT      87 a      162 c      166 g      122 t
ORIGIN
Query Match      3.9%; Score 67; DB 14; Length 537;
Best Local Similarity 100.0%; Pred. No. 4.9e-22;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      410 AGTGTGACACGAGAGGCTGGGGAGTCTCGACACCTGATCTGCGACCCAGCCA 469
Db      471 AGTGTGACACGAGAGGCTGGGGAGTCTCGACACCTGATCTGCGACCCAGCCA 530
Oy      470 GGGCCTG 476
Db      531 GGGCCTG 537

RESULT 4
LOCUS   BB849097
DEFINITION BB849097 RIKEN full-length enriched, adult inner ear Mus musculus
ACCESSION BB849097
VERSION   BB849097
KEYWORDS  424 bp mRNA linear EST 26-NOV-2001
SOURCE    cDNA clone F930006502 5', mRNA sequence.
ORGANISM  Mus musculus
REFERENCE 1 (bases 1 to 424)
AUTHORS   Akimura, T., Arakawa, T., Carninci, P., Furuno, M., Hanagaki, T.,
            Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Imotani, K., Ishii
            Y., Ito, M., Kawai, J., Kojima, Y., Kondo, H., Kouda, M., Matsuyama, T.,
            Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Okazaki, Y., Okido, T.,
            Saito, R., Sakai, C., Sakai, K., Sakazume, N., Sasaki, D., Sato, K.,
            Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagawa

```

```

VERSION   B0560868.1 GI:21461753
KEYWORDS  EST.
SOURCE    house mouse.
ORGANISM  Mus musculus
REFERENCE 1 (bases 1 to 537)
AUTHORS   Vanburen, V., Plao, Y., Dudekula, D.B., Qian, Y., Carter, M.G., Martin
            P.R., Stagg, C.A., Bassey, U., Alba, K., Hanatani, T., Katgul, G.J.,
            Luo, A.G. and Ko, M.S.H.
            Assembly, verification, and initial annotation of NIA 7.4K mouse
            cDNA clone set
            Unpublished (2002)
            Contact: Yong Qian
            Laboratory of Genetics
            National Institute on Aging/National Institutes of Health
            333 Cassell Drive, Suite 3000, Baltimore, MD 21224-6820, USA
            Email: cdna@igsun.grc.nia.nih.gov
            This clone set has been freely distributed to the community. Please
            visit http://lgsun.grc.nia.nih.gov/cDNA/NIA_7.4K.html for details.
            Plate: H4067 row: A column: 01
            Seq primer: -21M13 Reverse
            High quality sequence stop: 537
            POLYA=No.
            Location/Qualifiers
                1..537
                /organism="Mus musculus"
                /strain="C57BL/6"
                /db_xref="taeST:H4067A01-5"
                /db_xref="taxon:10090"
                /clone="H4067A01"
                /clone_lib="NIA Mouse 7.4K cDNA Clone Set"
                /sex="mixed"
                /dev_stage="mixed"
                /lab_host="DH10B"
                /note="Vector: pSPORT1; site_1: SalI; site_2: NotI; This
                clone is among a reartayed set of 7,407 clones from more
                than 20 cDNA libraries."

BASE COUNT      87 a      162 c      166 g      122 t
ORIGIN
Query Match      3.9%; Score 67; DB 14; Length 537;
Best Local Similarity 100.0%; Pred. No. 4.9e-22;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      410 AGTGTGACACGAGAGGCTGGGGAGTCTCGACACCTGATCTGCGACCCAGCCA 469
Db      471 AGTGTGACACGAGAGGCTGGGGAGTCTCGACACCTGATCTGCGACCCAGCCA 530
Oy      470 GGGCCTG 476
Db      531 GGGCCTG 537

RESULT 4
LOCUS   BB849097
DEFINITION BB849097 RIKEN full-length enriched, adult inner ear Mus musculus
ACCESSION BB849097
VERSION   BB849097
KEYWORDS  424 bp mRNA linear EST 26-NOV-2001
SOURCE    cDNA clone F930006502 5', mRNA sequence.
ORGANISM  Mus musculus
REFERENCE 1 (bases 1 to 424)
AUTHORS   Akimura, T., Arakawa, T., Carninci, P., Furuno, M., Hanagaki, T.,
            Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Imotani, K., Ishii
            Y., Ito, M., Kawai, J., Kojima, Y., Kondo, H., Kouda, M., Matsuyama, T.,
            Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Okazaki, Y., Okido, T.,
            Saito, R., Sakai, C., Sakai, K., Sakazume, N., Sasaki, D., Sato, K.,
            Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagawa

```

A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T.,  
 Watanabe, A., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.  
 RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al.  
 2001)  
 JOURNAL  
 COMMENT  
 Unpublished (2001)  
 Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-resgsc.riken.go.jp,  
 URL: http://genome.gsc.riken.go.jp/  
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh  
 M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new  
 genes. Genome Res. 10 (10), 1617-1630 (2000)  
 Wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,  
 Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura  
 S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and  
 Hayashizaki, Y.  
 RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
 10 (11), 1757-1771 (2000)  
 Kono, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara  
 Y. and Hayashizaki, Y.  
 Computer-based methods for the mouse full-length cDNA  
 encyclopedia: real-time sequence clustering for construction of a  
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for  
 further details.  
 e mouse tissues.  
 FEATURES  
 source  
 1. 424  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="F930006G02"  
 /clone\_1lb="RIKEN full-length enriched, adult inner ear"  
 /tissue\_type="inner ear"  
 /dev\_stage="adult"  
 /note="pooled tissues : (tissue\_type=cerebellum,  
 dev\_stage=16 days neonate, sex=mixed),  
 (tissue\_type=cerebellum, dev\_stage=0 day neonate,  
 sex=mixed), (tissue\_type=hippocampus, dev\_stage=adult,  
 sex=male), (tissue\_type=whole body, dev\_stage=9 days  
 embryo, sex=mixed), (tissue\_type=lung, dev\_stage=13 days  
 embryo, sex=mixed)"  
 BASE COUNT 65 a 124 c 132 g 103 t  
 ORIGIN  
 Query Match 2.9% Score 49; DB 10; Length 424;  
 Best Local Similarity 100.0%; Pred. No. 4.5e-13;  
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 410 ACTGTGTGCACGAGCTGGGAGCTGCTGCACCACTGCATGTCTGC 458  
 ||||||||||||||||||||||||||||||||||||||||||||  
 Db 373 ACTGTGTGCACGAGCTGGGAGCTGCTGCACCACTGCATGTCTGC 421  
 RESULT 5  
 LOCUS BB374499 315 bp mRNA linear EST 13-JUL-2000  
 DEFINITION BB374499 RIKEN full-length enriched, 16 days embryo head Mus  
 musculus cDNA clone C130075G16 3' similar to A126063 Mus musculus  
 connective tissue growth factor-like protein precursor (Ctgfl) mRNA  
 , mRNA sequence.  
 ACCESSION BB374499  
 VERSION BB374499.1 GI:9086993  
 KEYWORDS EST.  
 SOURCE house mouse.

ORGANISM  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE  
 AUTHORS  
 1 (bases 1 to 315)  
 Kono, H., Aizawa, K., Akahira, S., Akiyama, J., Arakawa, T., Carninci  
 P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N.,  
 Hirozane, T., Hori, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M.,  
 Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N.,  
 Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurihara, C., Kusabe, M.,  
 Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y.,  
 Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K., Shibata, K., Shibata  
 Y., Shigemoto, Y., Shinnaga, A., Shitaki, T., Sogabe, Y., Sugahara, Y.,  
 Suzuki, H., Suzuki, H., Tagawa, A., Takahashi, F., Tomioka, N., Toya  
 T., Tsunoda, Y., Watanabe, S., Watanabe, S., Yamamura, T., Yamana, I.,  
 Yano, R., Yasunishi, A., Yokota, T., Yoshida, K., Yoshiki, A., Yoshino  
 M., Muramatsu, M. and Hayashizaki, Y.  
 RIKEN Mouse ESTs (Kono, H., et al.)  
 JOURNAL  
 COMMENT  
 Unpublished (2000)  
 Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-resgsc.riken.go.jp,  
 URL: http://genome.gsc.riken.go.jp/  
 Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoka, S., Sasaki  
 N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 Thermostabilization and thermoactivation of thermostable enzymes by  
 trehalose and its application for the synthesis of full length  
 cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)  
 Itoh, M., Kitsuana, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J.,  
 Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki  
 Y. and Hayashizaki, Y.  
 Automated filtration-based high-throughput plasmid preparation  
 system. Genome Res. 9 (5), 463-470 (1999)  
 Carninci, P. and Hayashizaki, Y.  
 High-efficiency full-length cDNA cloning. Methods Enzymol. 303,  
 19-44 (1999)  
 Please visit our web site (<http://genome.riken.go.jp>) for  
 further details.  
 FEATURES  
 source  
 1. 315  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="C130075G16"  
 /clone\_1lb="RIKEN full-length enriched, 16 days embryo  
 head"  
 /sex="mixed"  
 /tissue\_type="head"  
 /dev\_stage="16 days embryo"  
 /lab\_host="DH10B"  
 /note="Site\_1: SalI; Site\_2: BamHI. cDNA library was  
 prepared and sequenced in Mouse Genome Encyclopedia  
 Project of Genome Exploration Research Group in Riken  
 Genomic Sciences Center and Genome Science Laboratory in  
 RIKEN. Division of Experimental Animal Research in Riken  
 contributed to prepare mouse tissues. 1st strand cDNA was  
 primed with a primer [5'-GAGAGAGAGATTCGAGTAAATATCCCGCCCCC 3']. cDNA was  
 prepared by using trehalose thermo-activated reverse  
 transcriptase and subsequently enriched for full-length by  
 cap-trapper. Second strand cDNA was prepared with the  
 primer adapter of sequence [5'-GAGAGAGATTCGAGTAAATATCCCGCCCCC 3']. cDNA  
 was cloned into the XhoI and BamHI sites. Vector: a  
 modified pBluescript KS(+) after bulk excision from Lambda  
 FLC 1"  
 BASE COUNT 87 a 80 c 67 g 81 t  
 ORIGIN

Query Match 2.0%; Score 35; DB 10; Length 315;  
 Best Local Similarity 100.0%; Pred. No. 4,1e-06;  
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1610 TCTGCAGCCAGGCGCTTCTCTCAGCATGAGA 1644  
 |||  
 207 TCTGCAGGCGCAGGCGCTTCTCTCAGCATGAGA 241

RESULT 6  
 BB792544 369 bp mRNA linear EST 16-NOV-2001  
 LOCUS BB792544 RIKEN full-length enriched, kidney CCL-142 RAG CDNA Mus  
 DEFINITION musculus CDNA clone G430141015 3', mRNA sequence.  
 BB792544  
 VERSION BB792544.1 GI:16961776  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 369)

REFERENCE  
 AUTHORS Akimura, T., Arakawa, T., Carninci, P., Furuno, M., Hanagaki, T.,  
 Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Imotani, K., Ishii,  
 Y., Ito, M., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Matsuyama, T.,  
 Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Okazaki, Y., Okido, T.,  
 Saito, R., Sakai, C., Sakai, K., Sakazume, N., Sasaki, D., Sato, K.,  
 Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagawa,  
 A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toyota, T.,  
 Watanishi, A., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.  
 RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al.  
 2001)

TITLE Unpublished (2001)

JOURNAL Contact: Yoshinori Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-res@gsr.riken.go.jp/  
 URL: http://genome.gsc.riken.go.jp/  
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh,  
 M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new  
 genes. Genome Res. 10 (10), 1617-1630 (2000)

Wag1, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,  
 Watanishi, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura,  
 S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and  
 Hayashizaki, Y.  
 RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
 10 (11), 1757-1771 (2000)

Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara,  
 Y. and Hayashizaki, Y.  
 Computer-based methods for the mouse full-length cDNA  
 encyclopedia: real-time sequence clustering for construction of a  
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
 Please visit our web site (http://genome.gsc.riken.go.jp) for  
 further details.  
 e mouse tissues.

FEATURES  
 source location/Qualifiers  
 1. 369  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /clone="G430141015"  
 /clone\_lib="RIKEN full-length enriched, kidney CCL-142 RAG  
 CDNA"  
 /tissue\_type="kidney"  
 /cell\_line="CCL-142 RAG"  
 /note="pooled cell lines : (cell\_line=CRL-1751 WEHI 164),  
 (cell\_line=CRL-2116 JC), (cell\_line=RCB-0035 WEHI-3),

Query Match 2.0%; Score 35; DB 10; Length 369;  
 Best Local Similarity 100.0%; Pred. No. 4,2e-06;  
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1610 TCTGCAGCCAGGCGCTTCTCTCAGCATGAGA 1644  
 |||  
 262 TCTGCAGGCGCAGGCGCTTCTCTCAGCATGAGA 296

RESULT 7  
 A1225477 418 bp mRNA linear EST 29-OCT-1998  
 LOCUS ue88b01.y1 Soares\_NMPu Mus musculus CDNA clone IMAGE:1498153 5',  
 DEFINITION mRNA sequence.  
 A1225477  
 VERSION A1225477.1 GI:3808530  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 418)

REFERENCE  
 AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,  
 Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,  
 Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,  
 Theising, B., Wyllie, T., Lennon, G., Soares, B., Wilson, R. and  
 Waterston, R.  
 The WashU-HMI Mouse EST Project  
 Unpublished (1996)  
 Contact: Marra M/Mouse EST Project  
 WashU-HMI Mouse EST Project  
 Washington University School of Medicine  
 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: mouseest@watson.wustl.edu  
 This clone is available royalty-free through LNL; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 MGI:935757  
 Seq primer: -40RP from Gibco  
 High quality sequence stop: 403.

FEATURES  
 source location/Qualifiers  
 1. 418  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:1498153"  
 /clone\_lib="Soares\_NMPu"  
 /sex="female"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /note="Organ: uterus; Vector: pTZ193D-Pac (Pharmacia) with  
 a modified polylinker; 1st strand cDNA was prepared from  
 pregnant mouse uterus, and was then primed with a Not I -  
 oligo(dT) primer. Double-stranded cDNA was ligated to Eco  
 RI adaptors (Pharmacia), digested with Not I and cloned  
 into the Not I and Eco RI sites of the modified pTZ193

vector. Library is normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo.

Query Match 2.0%; Score 35; DB 9; Length 418;  
Best Local Similarity 100.0%; Pred. No. 4.3e-06;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1610 TCTGCGAGCGAGCGCTTCTCTCAGCATGAGA 1644  
DB 300 TCTGCGAGCGAGCGCTTCTCTCAGCATGAGA 334

RESULT 8 488 bp mRNA linear EST 21-JAN-1998  
AA754979  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

The WashU-HMI Mouse EST Project  
Unpublished (1996)  
Contact: Maria M/Mouse EST Project  
WashU-HMI Mouse EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mouseest@wustl.edu  
This clone is available royalty-free through LLNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
MGI:642454  
Seq primer: -28ml3 rev2 ET from Amer sham  
High quality sequence stop: 478.  
Location/Qualifiers

FEATURES  
source

1. 488  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone\_id="IMAGE:1195358"  
/clone\_lib="Soares\_mammary\_gland\_NBMWG"  
/sex="male"  
/tissue\_type="mammary gland"  
/dev\_stage="4 weeks"  
/lab\_host="DH10B"  
/note="Organ: mammary gland; Vector: pTZ19-Pac (Pharmacia  
RI: 1st strand cDNA was primed with a Not I - oligo(dT)  
primer [5'  
TGTATACCATCTGAAGTGGAGCGCCGAGATGTTTTTTTTTTTTTTTTTTTT  
T 3']; double-stranded cDNA was ligated to Eco RI  
adaptors (Pharmacia), digested with Not I and cloned into  
the Not I and Eco RI sites of the modified pTZ19 vector.  
RNA provided by Dr. Minoru Ko, Wayne State Univ. Library  
constructed and normalized by Bento Soares and M. Fatima  
Bonaldo."

BASE COUNT 110 a 138 c 130 g 110 t

Query Match 2.0%; Score 35; DB 9; Length 488;  
Best Local Similarity 100.0%; Pred. No. 4.4e-06;

Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 897 TCCACGAGACGATTCGCAACGAGGATCCA 931  
DB 14 TCCACGAGACGATTCGCAACGAGATCCA 48

RESULT 9 792 bp mRNA linear EST 24-OCT-2000  
BF138093  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-remail.nih.gov  
Tissue Procurement: Gilbert Smith, Ph.D.  
DNA Library Preparation: Life Technologies, Inc.  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA sequencing by: Incyte Genomics, Inc.  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM9254 row: f column: 02  
High quality sequence stop: 705.  
Location/Qualifiers

FEATURES  
source

1. 792  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone\_id="IMAGE:4012801"  
/clone\_lib="NCI-CGAP-Lu30"  
/tissue\_type="tumor, metastatic to mammary"  
/lab\_host="DH10B"  
/note="Organ: lung; Vector: PCMV-SPORE6; Site\_1: NotI;  
Site\_2: SalI; transgenic model WNT-1, expression driven by  
MMTV-LTR enhancer; Cloned unidirectionally. Primer: Oligo  
dT. Library constructed by Life Technologies.  
Investigator providing samples: Gilbert Smith, NIH"

BASE COUNT 193 a 204 c 205 g 190 t

Query Match 2.0%; Score 35; DB 12; Length 792;  
Best Local Similarity 100.0%; Pred. No. 4.8e-06;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1610 TCTGCGAGCGAGCGCTTCTCTCAGCATGAGA 1644  
DB 659 TCTGCGAGCGAGCGCTTCTCTCAGCATGAGA 693

RESULT 10 966 bp mRNA linear EST 24-OCT-2000  
BF141695  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT





Query Match 1.9%; Score 32; DB 12; Length 380;  
 Best Local Similarity 100.0%; Pred. No. 0.00013;  
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 375 GTACCCCTGCTGCTGATGGCTGCTGCTG 406  
 ||||||||||||||||||||||||||||||||||  
 DB 206 GTACCCCTGCTGCTGATGGCTGCTGCTG 237

RESULT 13  
 BG900069 405 bp mRNA linear EST 06-NOV-2001  
 LOCUS HOA51-1-A11.R HOA (Human Osteoarthritic Cartilage) Homo sapiens  
 DEFINITION CDNA, mRNA sequence.  
 ACCESSION BG900069  
 VERSION BG900069.1 GI:14310318  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 Kumar, S., Connor, J.R., Dodds, R.A., Halsey, W., Van Horn, M., Mao, J.,  
 Sathie, G., Mul, P., Agarwal, P., Badger, A.M., Lee, J.C., Gowen, M. and  
 Larf, M.W.  
 Identification and initial characterization of 5000 expressed  
 sequenced tags (ESTs) each from adult human normal and  
 osteoarthritic cartilage cDNA libraries  
 Osteoarthr. Cartil. 9 (7), 641-653 (2001)  
 21482651  
 JOURNAL Contact: Sanjay Kumar  
 MEDLINE UW2109  
 COMMENT GlaxoSmithKline  
 709 Swedeland Road, P.O. Box 1539, King of Prussia, PA 19406, USA  
 Tel: 610-270-7245  
 Fax: 610-270-5598  
 Email: sanjay\_kumar-1@sk.com  
 Seq primer: 17.

FEATURES  
 source Location/Qualifiers  
 1..405  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="HOA (Human Osteoarthritic Cartilage)"  
 /tissue\_type="Cartilage"  
 /lab\_host="E.coli DH10 B"  
 /note="Vector: pSPORT I; Site\_1: SalI; Site\_2: NotI;  
 Directional"  
 BASE COUNT 62 a 140 c 135 g 68 t  
 ORIGIN

Query Match 1.9%; Score 32; DB 12; Length 405;  
 Best Local Similarity 100.0%; Pred. No. 0.00013;  
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 375 GTACCCCTGCTGCTGATGGCTGCTGCTG 406  
 ||||||||||||||||||||||||||||||||||  
 DB 202 GTACCCCTGCTGCTGATGGCTGCTGCTG 233

RESULT 14  
 BM751866 489 bp mRNA linear EST 04-MAR-2002  
 LOCUS K-EST0028180 S2SN068s1 Homo sapiens CDNA clone S2SN068s1-2-D09  
 DEFINITION 5', RNA sequence.  
 ACCESSION BM751866  
 VERSION BM751866.1 GI:19081484  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 489)

AUTHORS Kim, N.S., Bahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R.,  
 Oh, K.J., Cheong, J.E., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and  
 Kim, Y.S.  
 TITLE 21C Frontier Korean EST Project 2001.  
 JOURNAL Unpublished (2002)  
 COMMENT Contact: Kim YS  
 Genome Research Center  
 Korea Research Institute of Bioscience & Biotechnology  
 52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
 Tel: +82-42-860-4470  
 Fax: +82-42-860-4409  
 Email: yongsung@mail.kribb.re.kr  
 Plate: 2 row: D column: 09  
 High quality sequence stop: 489.

FEATURES  
 source Location/Qualifiers  
 1..489  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="S2SN068s1-2-D09"  
 /clone\_lib="S2SN068s1"  
 /sex="M"  
 /tissue\_type="Ascites"  
 /cell\_type="Epithelial"  
 /cell\_line="SNU-668"  
 /lab\_host="Top10F"  
 /note="Organ: Stomach; Vector: pCNS; Site\_1: EcoRI;  
 Site\_2: NotI; The poly (A)+ RNA was decapep with tabacco  
 acid pyrophosphatase (TAP) and ligated with DNA-RNA linker  
 including EcoRI site by treatment of T4 RNA ligase. The  
 first strand cDNA was synthesized from oligo dT-selected  
 mRNA by priming with dT-tailed vector. The dT-tailed  
 vector was adjusted to have about 60nt. The cDNA vector  
 was circularized with E. coli DNA ligase after digestion  
 of EcoRI which site is also included in vector. An RNA  
 strand converted to a DNA strand by Okayama-Berg method.  
 The obtained cDNA vectors were used for transformation of  
 competent cells E. coli Top10F by electroporation method.  
 After analyzing and sequencing about 2,000 ~ 3,000  
 colonies in original cDNA library, the abundant cDNAs were  
 selected and amplified by PCR reaction using vector region  
 primer including 77 promoter as 5' primer and N(dT)14 as  
 3' primer. The PCR products were used as template for  
 synthesis of biotinylated single stranded RNA by in vitro  
 transcription reaction. The synthesized RNA probes were  
 hybridized with antisense single stranded cDNAs prepared  
 from original library and incubated with avidin-gel.  
 After removing DNA-RNA hybrids by centrifuge, the  
 subtracted cDNA libraries were constructed by  
 transformation of the remaining DNA into competent cells E.  
 coli Top10F with electroporation method."  
 BASE COUNT 75 a 173 c 159 g 82 t  
 ORIGIN

Query Match 1.9%; Score 32; DB 14; Length 489;  
 Best Local Similarity 100.0%; Pred. No. 0.00014;  
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 654 GAGATGTCGGCTGCCCGAGTGGAGTGGCC 685  
 ||||||||||||||||||||||||||||||||||  
 DB 139 GAGATGTCGGCTGCCCGAGTGGAGTGGCC 170

RESULT 15  
 BG538695 618 bp mRNA linear EST 03-APR-2001  
 LOCUS 602566932F1 NIH\_MGC\_77 Homo sapiens CDNA clone IMAGE:4691574 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BG538695  
 VERSION BG538695.1 GI:13530928  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE  
1 (bases 1 to 618)  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Mammalia; Euthera; Primates; Catarrhini; Homiidae; Homo.  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
Tissue Procurement: CLONTECH Laboratories, Inc.  
CDNA Library Preparation by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: L1CM1510 row: P column: 07  
High quality sequence stop: 499.

FEATURES  
source  
Location/Qualifiers

1..618  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4691574"  
/lab\_host="NIH-MGC\_77"  
/lab\_host="DH10B (T1 phage-resistant)"  
/note="Organ: lung; Vector: pDNR-LIB (Clontech); Site\_1:  
SfiI (ggccgctcgcc); Site\_2: SfiI (ggccattatggcc); 5' and  
3' adaptors were used in cloning as follows: 5' adaptor  
sequence: 5'-CAGGCCCATTTATGGCC-3' and 3' adaptor sequence:  
5'-ATTCTAGAGCCGCGAGCGCCGACATG-dT(30)BN-3' (where B = A,  
C, or G and N = A, C, G, or T). Average insert size 1.9  
kb (range 0.5-4.0 kb). 12/15 colonies contained inserts  
by PCR. This library was enriched for full-length clones  
and was constructed by Clontech Laboratories (Palo Alto,  
CA). Note: this is a NIH-MGC Library."

BASE COUNT  
ORIGIN  
95 a 194 c 213 g 116 t

Query Match 1.9%; Score 32; DB 12; Length 618;  
Best Local Similarity 100.0%; Pred. No. 0.00014;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 375 GTACCCCTGCTGCTGATGCTGCTGCTGCTG 406  
DB 291 GTACCCCTGCTGCTGATGCTGCTGCTGCTG 322

Search completed: July 28, 2003, 18:02:02  
Job time : 3040.91 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 15:19:16 ; Search time 486.326 Seconds

(without alignments)  
10472.371 Million cell updates/sec

Title: US-10-010-408-1\_COPY\_1534\_1708

Perfect score: 175  
Sequence: 1 AGTCCAGCACTTGACCTT.....GCTAGATAAACACCCAAA 175

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 2054640 seqs, 14551402878 residues

Word size: 0

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

```
GenBank1:
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_ov:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rtd:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*
```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1          | 102   | 58.3        | 137964 | 2     | AC126895    |
| 2          | 102   | 58.3        | 226303 | 2     | AC095418    |
| 3          | 69    | 39.4        | 1741   | 10    | AF259981    |
| 4          | 35    | 20.0        | 1734   | 6     | AR210324    |
| 5          | 35    | 20.0        | 1734   | 6     | AR210325    |
| 6          | 35    | 20.0        | 1734   | 10    | AF100778    |
| 7          | 35    | 20.0        | 61072  | 10    | AL731698    |
| 8          | 35    | 20.0        | 216757 | 2     | AL669906    |
| 9          | 33    | 18.9        | 1739   | 10    | AF126063    |
| 10         | 23    | 13.1        | 145540 | 2     | AC015962    |
| 11         | 23    | 13.1        | 176107 | 2     | AP002393    |
| 12         | 23    | 13.1        | 176612 | 2     | AC090246    |
| 13         | 23    | 13.1        | 185568 | 9     | AC105227    |
| 14         | 21    | 12.0        | 167691 | 2     | AC103349    |
| 15         | 21    | 12.0        | 173702 | 2     | AC114696    |
| 16         | 21    | 12.0        | 231791 | 2     | AL732620    |
| 17         | 20    | 11.4        | 25546  | 2     | AC017886    |
| 18         | 20    | 11.4        | 49759  | 3     | AE002717    |
| 19         | 20    | 11.4        | 74534  | 3     | AC004735    |
| 20         | 20    | 11.4        | 102714 | 9     | AP000692    |
| 21         | 20    | 11.4        | 106735 | 3     | AC006215    |
| 22         | 20    | 11.4        | 127724 | 2     | AC129134    |
| 23         | 20    | 11.4        | 138653 | 9     | AL445431    |
| 24         | 20    | 11.4        | 148879 | 2     | AC115134    |
| 25         | 20    | 11.4        | 159119 | 2     | AC120233    |
| 26         | 20    | 11.4        | 160929 | 2     | AC008677    |
| 27         | 20    | 11.4        | 164064 | 2     | AC044809    |
| 28         | 20    | 11.4        | 175835 | 2     | AC024074    |
| 29         | 20    | 11.4        | 184681 | 2     | AC084409    |
| 30         | 20    | 11.4        | 189672 | 3     | CNS01DVM    |
| 31         | 20    | 11.4        | 191895 | 3     | AC093197    |
| 32         | 20    | 11.4        | 206803 | 2     | AC115932    |
| 33         | 20    | 11.4        | 309928 | 2     | AE003666    |
| 34         | 20    | 11.4        | 340000 | 9     | AP001725    |
| 35         | 19    | 10.9        | 60372  | 2     | AC097209    |
| 36         | 19    | 10.9        | 140726 | 2     | AC122991    |
| 37         | 19    | 10.9        | 146463 | 2     | AC068306    |
| 38         | 19    | 10.9        | 146937 | 2     | AC109876    |
| 39         | 19    | 10.9        | 160199 | 9     | AF391285    |
| 40         | 19    | 10.9        | 163713 | 2     | AC069432    |
| 41         | 19    | 10.9        | 173797 | 2     | AC027289    |
| 42         | 19    | 10.9        | 178980 | 10    | AL669897    |
| 43         | 19    | 10.9        | 181721 | 9     | AC010879    |
| 44         | 19    | 10.9        | 182522 | 9     | AC015689    |
| 45         | 19    | 10.9        | 189377 | 2     | AC113164    |

## ALIGNMENTS

RESULT 1  
AC126895  
LOCUS  
DEFINITION  
AC126895 137964 bp DNA linear HTG 24-JUL-2002  
Rattus norvegicus clone CH230-301E4, \*\*\* SEQUENCING IN PROGRESS  
\*\*\* 49 unordered pieces.

ACCESSION  
AC126895.1 GI:21724040  
VERSION  
HTG: HTGS-PHASE1.  
KEYWORDS  
Rattus norvegicus  
SOURCE  
Rattus norvegicus  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE  
1 (bases 1 to 137964)  
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,

Alsbrooks, S.L., Amaralunga, H.C., Are, J.R., Ayele, M., Banks, T.,  
 Barbieri, J., Benton, J., Bimaga, K., Blankenburg, K., Bonin, D.,  
 Bowler, J., Bowler, S., Brileva, M., Brown, E., Brown, M., Bryant, N.P.,  
 Buhay, C., Burch, P., Burkett, C., Burrell, K.J., Byrd, N.C.,  
 Carron, T.F., Carter, M., Cavazos, S.R., Chaddo, J., Chavez, D.,  
 Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,  
 Cleveland, C.D., Cox, C., Coyle, M.D., Dathorpe, S.R., David, R.,  
 Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,  
 Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,  
 Doulatwalie, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,  
 Earhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,  
 Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,  
 Gabali, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,  
 Gortelli, J.H., Guevara, M., Gunaratne, P., Hale, S., Hamilton, K.,  
 Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,  
 Hernandez, O., Hodgson, A., Hognes, M., Hollaway, C., Hollins, B.,  
 Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,  
 Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudan, S.,  
 Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,  
 Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,  
 Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Louised, H.,  
 Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,  
 Maheshwari, M., Mapa, P., Martin, R., Martindale, A., Martinez, E.,  
 Massey, E., Mathney, E., McLeod, M.P., Meador, M., Mel, G., Metzger, M.,  
 Miner, G., Miner, Z., Mitchell, T., Mohabaty, K., Morgan, M., Morris, S.,  
 Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,  
 Nguyen, N., Nickerson, E., Nwokkenwo, S., Ogih, M., Okwunnu, G.,  
 Orangunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,  
 Peters, L., Pickens, R., Primus, E., Pul, L., Quiles, M., Ren, Y.,  
 Rivas, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Saverly, G.,  
 Scherer, S., Scott, G., Shen, H., Shooshari, N., Sisson, I.,  
 Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H.,  
 Sutton, A., Svatek, A., Taber, P., Tamerisa, A., Tamerisa, K., Tang, H.,  
 Tansy, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,  
 Usmani, K., Vasquez, L., Vera, V., Villalon, B., Vinson, R., Wang, Q.,  
 Wang, S., Ward-Moore, S., Warren, R., Washington, C., Wallington, S.,  
 Williams, G., Williamson, A., Wlezyk, R., Woodson, S., Worley, K.,  
 Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,  
 Wainstock, G., and Gibbs, R.

Direct Submission  
 Unpublished  
 2 (bases 1 to 137964)  
 Worley, K.C.  
 Direct Submission  
 Submitted (10-JUL-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 137964)

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL

COMMENT

Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: hgsc-help@bcm.tmc.edu  
 Project Information  
 Center project name: GZHG  
 Center clone name: CH230-301E4  
 Summary Statistics  
 Sequencing vector: Plasmid  
 Chemistry: Dye-terminator Big Dye 100% of reads  
 Assembly program: Phrap; version 0.990329  
 Consensus quality: 89474 bases at least Q40  
 Consensus quality: 93422 bases at least Q30  
 Consensus quality: 96506 bases at least Q20

\* NOTE: Estimated insert size may differ from sequence length  
 \* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 49 contigs. The true order of the pieces

\* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

1 1056: contig of 1056 bp in length  
 1057 1156: gap of unknown length  
 1157 2335: contig of 1179 bp in length  
 2336 2435: gap of unknown length  
 2436 3442: contig of 1007 bp in length  
 3443 3542: gap of unknown length  
 3543 5081: contig of 1539 bp in length  
 5082 5181: gap of unknown length  
 5182 6425: contig of 1243 bp in length  
 6425 7814: gap of unknown length  
 7814 7914: contig of 1290 bp in length  
 7915 9153: gap of unknown length  
 9154 9253: contig of 1239 bp in length  
 9254 10445: gap of unknown length  
 10446 10545: contig of 1192 bp in length  
 10546 10546: gap of unknown length  
 10546 11771: contig of 1226 bp in length  
 11772 11871: gap of unknown length  
 11871 13240: contig of 1369 bp in length  
 13241 13340: gap of unknown length  
 13341 14359: contig of 1019 bp in length  
 14360 14459: gap of unknown length  
 14460 15651: contig of 1192 bp in length  
 15652 15751: gap of unknown length  
 15752 17494: contig of 1743 bp in length  
 17495 17594: gap of unknown length  
 17595 18679: contig of 1085 bp in length  
 18680 20681: gap of unknown length  
 20681 20781: contig of 1902 bp in length  
 20782 22118: gap of unknown length  
 22118 22218: contig of 1337 bp in length  
 22218 23578: gap of unknown length  
 23578 23678: contig of 1360 bp in length  
 23679 25423: gap of unknown length  
 25424 25523: contig of 1745 bp in length  
 25524 27808: gap of unknown length  
 27809 27908: contig of 2285 bp in length  
 27909 30272: gap of unknown length  
 30273 30372: contig of 2364 bp in length  
 30373 33091: gap of unknown length  
 33091 33191: contig of 2719 bp in length  
 33192 33192: gap of unknown length  
 33192 35777: gap of 2586 bp in length  
 35778 35877: gap of unknown length  
 35878 36908: contig of 1031 bp in length  
 36909 37009: gap of unknown length  
 37009 38930: gap of unknown length  
 38931 39030: contig of 1922 bp in length  
 39031 41210: gap of unknown length  
 41210 41310: contig of 2160 bp in length  
 41311 42937: gap of unknown length  
 42937 42938: contig of 1627 bp in length  
 42938 46120: gap of unknown length  
 46121 46220: contig of 3083 bp in length  
 46221 46221: gap of unknown length  
 46221 48217: gap of 1997 bp in length  
 48218 48317: gap of unknown length  
 48318 50433: contig of 2116 bp in length  
 50434 50533: gap of unknown length  
 50534 52880: gap of 2347 bp in length  
 52881 52980: gap of unknown length  
 52981 55194: contig of 2214 bp in length  
 55195 55294: gap of unknown length  
 55295 57482: contig of 2188 bp in length  
 57483 57582: gap of unknown length  
 57583 60781: contig of 3139 bp in length  
 60782 60881: gap of unknown length  
 60882 62599: contig of 1718 bp in length

```

* 62600 62699: gap of unknown length
* 62700 65203: contig of 2504 bp in length
* 65204 65303: gap of unknown length
* 65304 66914: contig of 1611 bp in length
* 66915 67014: gap of unknown length
* 67015 70839: contig of 3825 bp in length
* 70840 70940 75139: gap of unknown length
* 70940 75139: contig of 4200 bp in length
* 75140 75239: gap of unknown length
* 75240 78296: contig of 3057 bp in length
* 78297 78396: gap of unknown length
* 78397 83138: contig of 4742 bp in length
* 83139 83338: gap of unknown length
* 83239 88204: contig of 4966 bp in length
* 88205 92238: gap of unknown length
* 88305 92238: contig of 3934 bp in length
* 92239 92339: gap of unknown length
* 92339 97339: contig of 5001 bp in length
* 97340 97439: gap of unknown length
* 103335 103534: contig of 6095 bp in length
* 103635 107080: gap of unknown length
* 107081 107180: gap of unknown length
* 107181 115201: contig of 8021 bp in length
* 115202 121237: gap of unknown length
* 121238 121337: contig of 5936 bp in length
* 121338 127910: gap of unknown length
* 127911 128010: contig of 6573 bp in length
* 128011 137964: gap of unknown length
* 137964 137964: contig of 9954 bp in length.
Location/Qualifiers
1. 137964

```

```

FEATURES
source
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/clone="CH230-301B4"

```

```

BASE COUNT 31831 a 34027 c 35229 g 31979 t 4898 others
ORIGIN

```

```

Query Match 58.3% Score 102; DB 2; Length 137964;
Best Local Similarity 99.3%; Pred. No. 1e-49;
Matches 152; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 1 AGTCCAGAGACTGAGCTTTGATTATTTTTCAGGAATGCACATCTCTTACGACTCGCAAAAC 60
DB 44819 AGTCCAGAGACTGAGCTTTGATTATTTTTCAGGAATGCACATCTCTTACGACTCGCAAAAC 44878
OY 61 AGGAAGGCTCCAGACACTCTGCGAGCGGCGCTTCTCTTCAGATGAGAAAGCAAGG 120
DB 44879 AGGAAGGCTCCAGACACTCTGCGAGCGGCGCTTCTCTTCAGATGAGAAAGCAAGG 44938
OY 121 GACAGCAGAGTACTCTCTCTGAGAGACTAGTC 153
DB 44939 GACAGCAGAGTACTCTCTCTGAGAGACTAGTC 44971

```

```

RESULT 2
AC095418 226303 bp DNA linear HTG 11-JUL-2002
LOCUS Rattus norvegicus clone CH230-7C10, *** SEQUENCING IN PROGRESS ***,
DEFINITION 51 unordered pieces.
ACCESSION AC095418
VERSION AC095418.3 GI:21717893
KEYWORDS HTG; HTGS; PHASE1.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 226303)
Muzny D.M., Adams G.C., Adio-Oduola B., Ali-osman F.R., Allen C.,
Albrioks S.L., Amaralunge H.C., Are J.R., Ayale M., Banks T.,
Bardaria O., Benton J., Blinze K., Blankenburg K., Bonnin D.,
Bouck J., Bowie S., Brieva M., Brown E., Brown M., Bryant N.P.,

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

```

Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,  
 Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,  
 Chen, G., Chen, R., Chen, Z., Chowdhury, I., Christopoulos, C.,  
 Cleveland, C.D., Cox, C., Coyle, M.D., Datome, S.R., David, R.,  
 Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,  
 Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,  
 Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,  
 Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escoto, M.,  
 Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,  
 Gabisi, A., Gao, Y., Garcia, A., Garner, T., Garza, N., Gill, R.,  
 Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,  
 Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,  
 Hernandez, O., Hodgson, A., Hughes, M., Holloway, C., Hollins, B.,  
 Homsi, F., Howard, S., Huber, J., Huliy, S., Hume, J., Jackson, L.E.,  
 Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudh, S.,  
 Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,  
 Kravtsov, J., Kuresh, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,  
 Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Louised, H.,  
 Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,  
 Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,  
 Massey, E., Mawhinney, E., McLeod, M.P., Meador, M., Mel, G., Metzger, M.,  
 Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,  
 Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,  
 Nguyen, N., Nickerson, E., Nwokwenkwo, S., Ogum, M., Okunolu, G.,  
 Oreguine, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,  
 Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,  
 Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savary, G.,  
 Scherer, S., Scott, G., Shen, H., Shoostrati, N., Sisson, I.,  
 Sodergren, E., Soneike, T., Sparks, A., Stanley, H., Stone, H.,  
 Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,  
 Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,  
 Usmank, K., Vasquez, L., Vera, V., Villalob, D., Vinson, R., Wang, Q.,  
 Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,  
 Williams, G., Williamson, A., Wleciyk, R., Wooden, S., Worley, K.,  
 Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorilla, S., Nelson, D.,  
 Weinstein, G., and Gibbs, R.

Direct Submission  
 Unpublished  
 2 (bases 1 to 226303)  
 Worley, K.C.  
 Direct Submission  
 Submitted (11-SEP-2001) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 226303)  
 Worley, K.C.  
 Direct Submission  
 Submitted (11-JUL-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On Jul 10, 2002 this sequence version replaced gi:17941885.

----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
 ----- Project Information  
 Center project name: CH230-7C10  
 Center clone name: CH230-7C10  
 ----- Summary Statistics  
 Sequencing vector: Plasmid  
 Chemistry: Dye-terminator Big Dye: 100% of reads  
 Assembly program: Phrap: version 0.990329  
 Consensus quality: 185950 bases at least Q40  
 Consensus quality: 190362 bases at least Q20  
 Consensus quality: 193076 bases at least Q20  
 -----

\* NOTE: Estimated insert size may differ from sequence length  
 \* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_drift\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_drift_data.html)).  
 \* NOTE: This is a "working draft" sequence. It currently  
 \* consists of 51 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as

```

* runs of N, but the exact sizes of the gaps are unknown.
* this record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
1051: contig of 1051 bp in length
1052 1151: gap of unknown length
1152 2286: contig of 1135 bp in length
2287 2386: gap of unknown length
2387 3433: contig of 1047 bp in length
3434 3533: gap of unknown length
3534 4614: contig of 1081 bp in length
4615 4714: gap of unknown length
4715 6581: contig of 1867 bp in length
6582 6681: gap of unknown length
6682 7773: contig of 1092 bp in length
7774 9530: contig of 1657 bp in length
9531 9630: gap of unknown length
9631 10651: contig of 1031 bp in length
10652 10761: gap of unknown length
10762 11958: contig of 1197 bp in length
11959 12058: gap of unknown length
12059 13140: contig of 1082 bp in length
13141 13240: gap of unknown length
13241 14777: contig of 1537 bp in length
14778 14877: gap of unknown length
14878 16063: contig of 1186 bp in length
16064 16163: gap of unknown length
16164 17585: contig of 1422 bp in length
17586 17685: gap of unknown length
17686 18916: contig of 1231 bp in length
18917 19016: gap of unknown length
19017 20305: contig of 1289 bp in length
20306 20405: gap of unknown length
20406 21537: contig of 1132 bp in length
21538 21637: gap of unknown length
21638 23703: contig of 2066 bp in length
23704 23803: gap of unknown length
23804 26306: contig of 2503 bp in length
26307 26406: gap of unknown length
26407 28431: contig of 2025 bp in length
28432 28531: gap of unknown length
28532 30615: contig of 2084 bp in length
30616 30715: gap of unknown length
30716 33215: contig of 2500 bp in length
33216 33315: gap of unknown length
33316 35357: contig of 2042 bp in length
35358 35457: gap of unknown length
35458 38012: contig of 2555 bp in length
38013 38112: gap of unknown length
38113 40338: contig of 2226 bp in length
40339 40438: gap of unknown length
40439 44051: contig of 3613 bp in length
44052 44151: gap of unknown length
44152 46294: contig of 2143 bp in length
46295 46394: gap of unknown length
46395 48845: contig of 2451 bp in length
48846 48945: gap of unknown length
48946 51753: contig of 2808 bp in length
51754 51853: gap of unknown length
51854 55915: contig of 4062 bp in length
55916 56015: gap of unknown length
56016 60468: contig of 4453 bp in length
60469 60568: gap of unknown length
60569 63640: contig of 3072 bp in length
63641 63740: gap of unknown length
63741 67888: contig of 4148 bp in length
67889 67988: gap of unknown length
67989 73014: contig of 5026 bp in length
73015 73114: gap of unknown length
73115 77413: contig of 4199 bp in length
77414 83786: gap of unknown length
83786: contig of 6373 bp in length

```

```

* 83787 83886: gap of unknown length
* 83887 89590: contig of 5704 bp in length
* 89591 89691: gap of unknown length
* 89691 93846: contig of 4155 bp in length
* 93847 93946: gap of unknown length
* 93947 98126: contig of 4180 bp in length
* 98127 98226: gap of unknown length
* 98227 104258: contig of 6032 bp in length
* 104259 104358: gap of unknown length
* 104359 110994: contig of 6636 bp in length
* 110995 111094: gap of unknown length
* 111095 119838: contig of 8744 bp in length
* 119839 119938: gap of unknown length
* 119939 126339: contig of 6401 bp in length
* 126340 126439: gap of unknown length
* 126440 134888: contig of 8449 bp in length
* 134889 134988: gap of unknown length
* 134989 142709: contig of 7721 bp in length
* 142710 142809: gap of unknown length
* 142810 149818: contig of 7009 bp in length
* 149819 149918: gap of unknown length
* 149919 158719: contig of 8801 bp in length
* 158720 158819: gap of unknown length
* 158820 169868: contig of 11049 bp in length
* 169869 169968: gap of unknown length
* 169969 180390: contig of 10422 bp in length
* 180391 180490: gap of unknown length
* 180491 191272: contig of 10782 bp in length
* 191273 191373: gap of unknown length
* 191373 209667: contig of 18295 bp in length
* 209668 209767: gap of unknown length
* 209768 226303: contig of 16536 bp in length.
* Location/Qualifiers
* 1.226303
FEATURES
source
BASE COUNT 57875 a 53533 c 52917 g 57041 t 5117 others
Query Match 58.3%; Score 102; DB 2; Length 226303;
Best Local Similarity 99.3%; Pred. No. 1e-49;
Matches 152; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 AGTCAGAGACTTGACTTGATTTTCAGAGATGCATCTCTTAAGCATCGCAAAAC 60
Db 113012 AGTCAGAGACTTGACTTGATTTTCAGAGATGCATCTCTTAAGCATCGCAAAAC 113071
QY 61 AGGAGAGCTCCACACTCTGCGAGCGGAGGCTTCTCTTACGATGAGAAAGCAAGG 120
Db 113072 AGGAGAGCTCCACACTCTGCGAGCGGAGGCTTCTCTTACGATGAGAAAGCAAGG 113131
QY 121 GACAGCAGAGTACTCTCTCTGAGGAGCTAAGTC 153
Db 113132 GACAGCAGAGTACTCTCTCTGAGGAGCTAAGTC 113164
RESULT 3
AF259981 1741 bp mRNA linear ROD 09-MAY-2000
LOCUS Rattus norvegicus CCN family protein COP-1 (COP-1) mRNA, complete
DEFINITION
cds:
AF259981
VERSION AF259981.1 GI:7739780
KEYWORDS
SOURCE Rattus norvegicus.
ORGANISM Rattus norvegicus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 1741)
Zhang,R., Averboukh,L., Zhu,W., Zhang,H., Jo,H., Dempsey,P.J.,
Coffey,R.J., Pardee,A.B. and Liang,P.
Identification of rCOP-1, a new member of the CCN protein family,

```



JOURNAL as a negative regulator for cell transformation  
MOL. CELL. BIOL. 18 (10): 6131-6141 (1998)

LOCUS 98414629

PUBMED 9742130

REFERENCE 2 (bases 1 to 1741)

AUTHORS Liang, P.

JOURNAL Direct Submission

Submitted (24-APR-2000) Cell Biology, Vanderbilt-Ingram Cancer Center, 649 MRB II, Nashville, TN 37232, USA

FEATURES Location/Qualifiers

1..1741

/organism="Rattus norvegicus"

/db\_xref="taxon:10116"

1..1741

/gene="Cop-1"

262..1014

/gene="Cop-1"

/note="secreted protein"

/product="CCN family protein COP-1"

/protein\_id="AAFG9011.1"

/db\_xref="GI:7739781"

/translation="MRGSPRLRLATSFLLULSNVCAQLCRPTCTPWPQCPQGV  
LVLDGCGCCVCAARLTSCSEHLVCEPSGLVCGPAGGAGAVCLDDEDCV  
NGRNYLDGERFKNCRVLCRDGDFCLPLCSDDVTLPSMDGFRPRRIQVPGKCE  
WVCDQGVTPAIDRSAAQGHOLSLVTPASADAPWPMNSTAMGPCSTTGLIATRVSN  
ONRRCOEIQRILCLPRCLARSHSSNSAF"

CDS

gene

source

BASE COUNT 386 a 491 c 480 g 384 t

ORIGIN

Query Match 39.4%; Score 69; DB 10; Length 1741;  
Best Local Similarity 100.0%; Pred. No. 6e-30;

Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 CTGAGCTTGATTTTCAGAGATGCACATCTCTTAAGCACTCCCAAAACAGAGAGCTC 70

DB 1557 CTGAGCTTGATTTTCAGAGATGCACATCTCTTAAGCACTCCCAAAACAGAGAGCTC 1616

QY 71 CACACCTCT 79

DB 1617 CACACCTCT 1625

RESULT 4 AR210324 1734 bp DNA linear PAT 20-JUN-2002  
LOCUS AR210324  
DEFINITION Sequence 17 from patent US 6387657.  
ACCESSION AR210324  
VERSION AR210324.1 GI:21512525  
KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1734)

AUTHORS Botstein, D.A., Cohen, R.L., Goddard, A.D., Gurney, A.L., Hillan, K.J.,  
Lawrence, D.A., Levine, A.J., Pennica, D., Roy, M., Ann. and Wood, W.I.

JOURNAL WISP polypeptides and nucleic acids encoding same

Patent: US 6387657-A 17 14-MAY-2002;

FEATURES Location/Qualifiers

1..1734

/organism="unknown"

BASE COUNT 355 a 491 c 495 g 393 t

ORIGIN

Query Match 20.0%; Score 35; DB 6; Length 1734;  
Best Local Similarity 100.0%; Pred. No. 1.5e-09;

Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 TCTGACAGCCAGGCGCTTCTCTTCAGCATGAGA 111

DB 1611 TCTGACAGCCAGGCGCTTCTCTTCAGCATGAGA 1645

RESULT 5 AR210325 1734 bp DNA linear PAT 20-JUN-2002  
LOCUS AR210325  
DEFINITION Sequence 18 from patent US 6387657.  
ACCESSION AR210325  
VERSION AR210325.1 GI:21512526  
KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1734)

AUTHORS Botstein, D.A., Cohen, R.L., Goddard, A.D., Gurney, A.L., Hillan, K.J.,  
Lawrence, D.A., Levine, A.J., Pennica, D., Roy, M., Ann. and Wood, W.I.

JOURNAL WISP polypeptides and nucleic acids encoding same

Patent: US 6387657-A 18 14-MAY-2002;

FEATURES Location/Qualifiers

1..1734

/organism="unknown"

BASE COUNT 393 a 495 c 491 g 355 t

ORIGIN

Query Match 20.0%; Score 35; DB 6; Length 1734;  
Best Local Similarity 100.0%; Pred. No. 1.5e-09;

Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 TCTGACAGCCAGGCGCTTCTCTTCAGCATGAGA 111

DB 124 TCTGACAGCCAGGCGCTTCTCTTCAGCATGAGA 90

RESULT 6 AF100778 1734 bp mRNA linear ROD 17-DEC-1998  
LOCUS AF100778  
DEFINITION Mus musculus connective tissue growth factor related protein WISP-2  
(WISP2) mRNA, complete cds.  
ACCESSION AF100778  
VERSION AF100778.1 GI:4028578  
KEYWORDS

SOURCE Mus musculus.

ORGANISM Mus musculus.

REFERENCE 1 (bases 1 to 1734)

AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

Pennica, D., Swanson, T.A., Welsh, J.W., Roy, M.A., Lawrence, D.A.,  
Lee, J., Brush, J., Taneyhill, L.A., Denel, B., Lew, M., Matanabe, C.,  
Cohen, R.L., Melhem, M.F., Finley, G.G., Quirk, P., Goddard, A.D.,  
Hillan, K.J., Gurney, A.L., Botstein, D. and Levine, A.J.

JOURNAL WISP genes are members of the connective tissue growth factor  
family that are up-regulated in wnt-1-transformed cells and  
aberrantly expressed in human colon tumors

Proc. Natl. Acad. Sci. U.S.A. 95 (25), 14717-14722 (1998)

FEATURES Location/Qualifiers

1..1734

/organism="Mus musculus"

/db\_xref="taxon:10090"

/cell\_line="C57MG"

/cell\_type="epithelial"

/tissue\_type="mammary"

/note="transformed by Wnt-1"

1..1734

/gene="Wisp2"

257..1012

/codon\_start=1

/product="connective tissue growth factor related protein  
WISP-2"

```

/protein_id="AAC96320.1"
/db_xref="GI:4028579"
/translation="MGNPLIHLAIFSLTSMYXSLCPAPCAPCPMTPOCPGCV
LVLDGGCCRCVARRLGGCDLHVCDGGLVCGGAPSGAGNCLPEEDGSCVE
NGRRVLDGTEFKPNCRRVLCRCDDGGCTCLPSCSEVDRLDSMCPRRRIQVGRCPCE
WYCDQAVMDPAIOPSSAQGHQLSALVTPASADGCPGPNMSTAMPGCPSTTGGLIATRV
NONRCQLEIQRICLSRPCILSRSHGSNNAF"
BASE COUNT      355 a      491 c      495 g      393 t
ORIGIN
Query Match      20.0%; Score 35; DB 10; Length 1734;
Best Local Similarity 100.0%; Pred.No. 1.5e-09;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      77 TCTGGCAGGCGCAGGCGCTTCTCTTCAGCATGAGA 111
|||||
Db      1611 TCTGGCAGGCGCAGGCGCTTCTCTTCAGCATGAGA 1645

RESULT 7
AL731698      61072 bp      DNA      linear      ROD 24-MAY-2002
LOCUS      Mouse DNA sequence from clone RP23-161B3 on chromosome 2, complete
DEFINITION
ACCESSION      AL731698
VERSION      AL731698.10 GI:21214309
KEYWORDS      HTG.
SOURCE      house mouse.
ORGANISM      Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 61072)
AUTHORS      Wallis, J.
TITLE      Direct Submission
JOURNAL      Submitted (24-MAY-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
On May 25, 2002 this sequence version replaced gi:21213601.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em: EMBL; Sw:
SWISSPROT; Tr: TrEMBL; Wp: WormPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-161B3 is
from the RPI-23 Mouse PAC Library
constructed by the group of Pieter de Jong.
For further details see http://www.chori.org/bacpac/home.htm
VECTOR: pBAC3.6.
FEATURES
source      1..61072
              Location/Qualifiers
              /organism="Mus musculus"
              /db_xref="taxon:10090"
              /chromosome="2"
              /clone="RP23-161B3"
              /clone_1fb="RP23-161B3"
              /clone_2fb="RP23-161B3"
BASE COUNT      15405 a      15368 c      15033 g      15266 t
ORIGIN
Query Match      20.0%; Score 35; DB 10; Length 61072;
Best Local Similarity 100.0%; Pred.No. 1.6e-09;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY      77 TCTGGCAGGCGCAGGCGCTTCTCTTCAGCATGAGA 111
|||||
Db      17823 TCTGGCAGGCGCAGGCGCTTCTCTTCAGCATGAGA 17857

RESULT 8
AL669906      216757 bp      DNA      linear      HTG 24-JUL-2002
LOCUS      Mus musculus chromosome 2 clone RP23-21762, *** SEQUENCING IN
DEFINITION
ACCESSION      AL669906
VERSION      AL669906.5 GI:21955520
KEYWORDS      HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE      Mus musculus.
ORGANISM      Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 216757)
AUTHORS      Sims, S.
TITLE      Direct Submission
JOURNAL      Submitted (23-JUL-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
On Jul 25, 2002 this sequence version replaced gi:18181793.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Web site: http://www.sanger.ac.uk
Contact: humquerry@sanger.ac.uk
----- Project Information
Center project name: bm217c2
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 211865 bases at least Q40
Consensus quality: 213403 bases at least Q30
Consensus quality: 214139 bases at least Q20
Insert size: 214757; sum-of-ctrls
Insert size: 234243; 1.88 error; agarose-fp
Quality coverage: 6.29x in Q20 bases; sum-of-ctrls Quality
coverage: 5.90x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 21 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 4000: contig of 4000 bp in length
* 4001 4100: gap of 100 bp
* 4101 15267: contig of 11167 bp in length
* 15268 15367: gap of 100 bp
* 15368 34970: contig of 19603 bp in length
* 34971 35070: gap of 100 bp
* 35071 39671: contig of 4601 bp in length
* 39672 39771: gap of 100 bp
* 39772 46360: contig of 6589 bp in length
* 46361 46460: gap of 100 bp
* 46461 49809: contig of 3349 bp in length
* 49810 49909: gap of 100 bp
* 49910 54509: contig of 4600 bp in length
* 54510 54609: gap of 100 bp
* 54610 65989: contig of 11380 bp in length
* 65990 66089: gap of 100 bp
* 66090 84635: contig of 18546 bp in length
* 84636 84735: gap of 100 bp
* 84736 87399: contig of 2664 bp in length
* 87400 87499: gap of 100 bp
* 87500 98601: contig of 11102 bp in length
* 98602 98701: gap of 100 bp

```

```

* 98702 105522: contig of 6821 bp in length
* 105523 105622: gap of 100 bp
* 105623 119773: contig of 14151 bp in length
* 119774 119873: gap of 100 bp
* 119874 134552: contig of 14679 bp in length
* 134553 134652: gap of 100 bp
* 134653 141202: contig of 6550 bp in length
* 141203 141302: gap of 100 bp
* 141303 145808: contig of 4506 bp in length
* 145809 145908: gap of 100 bp
* 145909 151572: contig of 5664 bp in length
* 151573 151672: gap of 100 bp
* 151673 169686: contig of 18014 bp in length
* 169687 169786: gap of 100 bp
* 169787 200095: contig of 30309 bp in length
* 200096 200195: gap of 100 bp
* 200196 213877: contig of 13882 bp in length
* 213878 213977: gap of 100 bp
* 213978 216757: contig of 2780 bp in length.
Location/Qualifiers

```

## FEATURES

```
1..216757
```

```
/organism="Mus musculus"
/db_xref="taxon:10090"

```

```
/chromosome="2"
/clone="RP23-217C2"

```

```
/clone_lib="RPC1-23"
1..4000

```

```
/note="assembly_fragment:03429
clone_end:sp6
vector_side:left"

```

```
misc_feature
15368..34970
/note="assembly_fragment:00814
fragment_chain:1"

```

```
misc_feature
35071..39671
/note="assembly_fragment:03544
fragment_chain:1"

```

```
misc_feature
39772..46360
/note="assembly_fragment:01438
fragment_chain:1"

```

```
misc_feature
46461..49809
/note="assembly_fragment:03410
fragment_chain:2"

```

```
misc_feature
49910..54509
/note="assembly_fragment:00811
fragment_chain:2"

```

```
misc_feature
54610..65989
/note="assembly_fragment:02217
fragment_chain:2"

```

```
misc_feature
66090..84635
/note="assembly_fragment:03529
fragment_chain:2"

```

```
misc_feature
84736..87399
/note="assembly_fragment:01019
fragment_chain:3"

```

```
misc_feature
87500..98601
/note="assembly_fragment:01170
fragment_chain:3"

```

```
misc_feature
98702..105522
/note="assembly_fragment:02829
fragment_chain:3"

```

```
misc_feature
105623..119773
/note="assembly_fragment:01476
fragment_chain:4"

```

```
misc_feature
119874..134552
/note="assembly_fragment:00505
fragment_chain:4"

```

```
misc_feature
134653..141202
/note="assembly_fragment:01951
fragment_chain:4"

```

```
misc_feature
141303..145808

```

```

/note="assembly_fragment:02637
fragment_chain:5"
145809..151572
/note="assembly_fragment:02352
fragment_chain:5"
151673..169686
/note="assembly_fragment:02207
fragment_chain:5"
169787..200095
/note="assembly_fragment:02849
fragment_chain:6"
200196..213877
/note="assembly_fragment:01052
fragment_chain:6"
213978..216757
/note="assembly_fragment:01318
clone_end:t7
vector_side:right"
BASE COUNT 56425 a 51573 c 51904 g 54852 t 2003 others
ORIGIN

```

```

Query Match 20.0%; Score 35; DB 2; Length 216757;
Best Local Similarity 100.0%; Pred. No. 1.7e-09;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```
OY 77 TGTGGAGCCAGCGCTTCTCTCAGCATGAGA 111
Db 88929 TGTGGAGCCAGCGCTTCTCTCAGCATGAGA ,88963
```

```

RESULT 9 1739 bp mRNA linear ROD 12-OCT-1999
AF126063
LOCUS
DEFINITION
(Mus musculus connective tissue growth factor-like protein precursor
(Ctgfl) mRNA, complete cds.
ACCESSION
AF126063
VERSION
AF126063.1 GI:4337059
KEYWORDS
SOURCE

```

```

ORGANISM
Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1739)
REFERENCE
Kumar S., Hand A.T., Connor J.R., Dadds R.A., Ryan P.J.,
Trill J.J., Fisher S.M., Nuttall M.E., Lipschutz D.B., Zou C.,
Hwang S.M., Volta B.J., James I.E., Riegan D.J., Gowen M. and
Lee J.C.

```

```

TITLE
Identification and cloning of a connective tissue growth
factor-like cDNA from human osteoblasts encoding a novel regulator
of osteoblast functions
JOURNAL
U. Biol. Chem. 274 (24), 17123-17131 (1999)
MEDLINE
99287915
PUBMED
10358067

```

```

REFERENCE
2 (bases 1 to 1739)
Kumar S. and Zou C.
TITLE
Direct Submission
JOURNAL
Submitted (04-FEB-1999) Bone & Cartilage Biology, UM 2109,
SmithKline Beecham, 709 Swedeland Rd., King of Prussia, PA 19406,
USA

```

```

FEATURES
1..1739
Location/Qualifiers

```

```

source
/organism="Mus musculus"
/db_xref="taxon:10090"

```

```

gene
/ctgfl_type="lung"
1..1739

```

```

CDS
/gene="Ctgfl"
242..997

```

```

/note="similar to the Mus musculus WISP-2 protein encoded
by the sequence presented in GenBank Accession Number
AF100778; putative growth factor; CTGF-L; contains IGF
binding (IGFBP), Von Willebrand factor type C (VWC) repeat
and thrombospondin type I (TSP1) domains; member of the

```

CCN (CTGF/CyR61/Nov) family; lacks the fourth carboxy-terminal (CT) domain present in other members of the CCN family"

/codon\_start=1

/product="connective tissue growth factor-like protein precursor"

/protein\_id="MAD18058.1"

/db\_xref="GI:4337060"

/translation="MRGNPLHLHLAISFLCISLWVYSQICPAPACAPMTPOCPGCV

LVLDDGCGCCVACARLRGSCDHLHVCDSQGLVCGPAGSGRGAACFEDDSCSEV

NCRRYLDGETFKPNCRCDDGCTGCLPSCEVRLPSMPCRRPRRLOVGRCPCE

WYCDVAVPAIOPSSAGOHLSALVTPASADGPCMSTANGPCSTTGLGIATFVS

NONRCOLEIQRICLRICSRCLSRSGSNMNAE"

BASE COUNT

375 a 480 c 489 g 395 t

ORIGIN

Query Match 18.9%; Score 33; DB 10; Length 1739;

Best Local Similarity 100.0%; Pred. No. 2.4e-08;

Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 79 TGGCAGCGCCAGGCGCTTCTCTTCAGCATGAGA 111

Db 1598 TGGCAGCGCCAGGCGCTTCTCTTCAGCATGAGA 1630

RESULT 10

AC015962

LOCUS Homo sapiens chromosome 18 clone RP11-111D6 map 18, WORKING DRAFT

DEFINITION SEQUENCE. 38 unordered pieces.

AC015962

AC015962.4 GI:7382674

HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULLTOP.

KEYWORDS

SOURCE

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

1 (bases 1 to 145540)

2 (bases 1 to 145540)

REFERENCE

AUTHORS

Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,

Baldwin, J., Barna, N., Beckerly, R., Boguski, L., Bouckgeater, B.,

Brown, A., Castle, A., Colangelo, M., Collins, S., Collymore, A.,

Cooke, P., Dearlano, K., Dewar, K., Doming, M., Donnell, L., Doyle, M.,

Ferreira, P., Fitzhugh, W., Forrest, C., Funke, R., Gage, D.,

Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heath, A., Horton, L.,

Howard, J. C., Johnson, R., Jones, C., Kann, L., Karas, A., Klein, J.,

Lehoczky, J., Liu, C., Locke, K., MacDonald, P., Margulis, N.,

McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrum, J.,

Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,

Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,

Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,

Testa, S., Tittrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,

Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.

Direct Submission

Submitted (17-NOV-1999) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

On Apr 1, 2000 this sequence version replaced gi:6957735.

ALL repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RV/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: MIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence\_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L970

Center clone name: 111\_D\_6

----- Summary Statistics

Sequencing vector: M13; M7815; 100% of reads

Chemistry: Dye-primer-amerham; 4% of reads  
Chemistry: Dye-terminator Big Dye; 96% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 125662 bases at least Q40  
Consensus quality: 134718 bases at least Q30  
Consensus quality: 138727 bases at least Q20  
Insert size: 80000; agarose-1p  
Insert size: 141840; sum-of-ctnigs  
Quality coverage: 5.5 in Q20 bases;  
Quality coverage: 3.1 in Q20 bases.  
NOTE: This is a 'working draft' sequence. It currently  
consists of 38 contigs. The true order of the pieces  
is not known and their order in this sequence record is  
arbitrary. Gaps between the contigs are represented as  
runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 1061 1160: contig of 1060 bp in length  
1061 1160: gap of 100 bp  
1161 2171: contig of 1011 bp in length  
2172 2271: gap of 100 bp  
2272 3611: contig of 1340 bp in length  
3612 3711: gap of 100 bp  
3712 5307: contig of 1536 bp in length  
5308 5407: gap of 100 bp  
5408 7422: contig of 2015 bp in length  
7423 7522: gap of 100 bp  
7523 9048: contig of 1526 bp in length  
9049 9148: gap of 100 bp  
9149 10173: contig of 1025 bp in length  
10174 10273: gap of 100 bp  
10274 11773: contig of 1500 bp in length  
11774 11873: gap of 100 bp  
11874 13476: contig of 1603 bp in length  
13477 13576: gap of 100 bp  
13577 15396: contig of 1820 bp in length  
15397 15496: gap of 100 bp  
15497 16981: contig of 1485 bp in length  
16982 17081: gap of 100 bp  
17082 19160: contig of 2079 bp in length  
19161 19260: gap of 100 bp  
19261 21188: contig of 1928 bp in length  
21189 21288: gap of 100 bp  
21289 23565: contig of 2277 bp in length  
23566 23665: gap of 100 bp  
23666 25713: contig of 2048 bp in length  
25714 25813: gap of 100 bp  
25814 26916: contig of 1103 bp in length  
26917 27016: gap of 100 bp  
27017 29802: contig of 2786 bp in length  
29803 29902: gap of 100 bp  
29903 32825: contig of 2923 bp in length  
32826 32925: gap of 100 bp  
32926 35752: contig of 2827 bp in length  
35753 35852: gap of 100 bp  
35853 37965: contig of 2113 bp in length  
37966 38065: gap of 100 bp  
38066 41361: contig of 3296 bp in length  
41362 41461: gap of 100 bp  
41462 44463: contig of 3002 bp in length  
44464 44563: gap of 100 bp  
44564 47687: contig of 3124 bp in length  
47688 47787: gap of 100 bp  
47788 51777: contig of 3990 bp in length  
51778 51877: gap of 100 bp  
51878 56131: contig of 4254 bp in length  
56132 56231: gap of 100 bp  
56232 60894: contig of 4663 bp in length  
60895 60994: gap of 100 bp  
60995 65362: contig of 4368 bp in length  
65363 65462: gap of 100 bp  
65463 69104: contig of 3642 bp in length

```

* 69105 69204: gap of 100 bp
* 69205 73085: contig of 3881 bp in length
* 73086 73185: gap of 100 bp
* 73186 77580: contig of 4395 bp in length
* 77581 77680: gap of 100 bp
* 77681 85326: contig of 7646 bp in length
* 85327 85426: gap of 100 bp
* 85427 92750: contig of 7324 bp in length
* 92751 92850: gap of 100 bp
* 92851 100690: contig of 7840 bp in length
* 100691 100790: gap of 100 bp
* 100791 108177: contig of 7387 bp in length
* 108178 108277: gap of 100 bp
* 108278 115728: contig of 7451 bp in length
* 115729 115828: gap of 100 bp
* 115829 125137: contig of 9309 bp in length
* 125138 125237: gap of 100 bp
* 125238 134119: contig of 8882 bp in length
* 134120 134219: gap of 100 bp
* 134220 145540: contig of 11321 bp in length.

```

```

FEATURES
source
1..145540 Homo sapiens
/db_xref="taxon:9606"
/chromosome="18"
/map="18"
/clone_lib="RP11-111D6"
/clone_lib="RP11-111D6"
1..1060

```

```

misc_feature
1..1060
/note="assembly-fragment"
1161..2171
/note="assembly-fragment"
2272..3611
/note="assembly-fragment"
3712..5307
/note="assembly-fragment"
5408..7422
/note="assembly-fragment"
7523..9048
/note="assembly-fragment"
9149..10173
/note="assembly-fragment"
10274..11773
/note="assembly-fragment"
11874..13476
/note="assembly-fragment"
13577..15396
/note="assembly-fragment"
15497..16981
/note="assembly-fragment"
17082..19160
/note="assembly-fragment"
19261..21188
/note="assembly-fragment"
21289..23565
/note="assembly-fragment"
23666..25713
/note="assembly-fragment"
25814..26916
/note="assembly-fragment"
clone_end:SP6
vector_side:left"
27017..29802
/note="assembly-fragment"
29903..32825
/note="assembly-fragment"
32926..35752
/note="assembly-fragment"
35853..37965
/note="assembly-fragment"
38066..41361
/note="assembly-fragment"
41462..44463
/note="assembly-fragment"

```

```

misc_feature /note="assembly-fragment"
44564..47687
/note="assembly-fragment"
misc_feature 47788..51777
/note="assembly-fragment"
misc_feature 51878..56131
/note="assembly-fragment"
misc_feature 56232..60894
/note="assembly-fragment"

Query Match 13.1%; Score 23; DB 2; Length 145540;
Best Local Similarity 100.0%; Pred. No. 0.027;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 11 CTTGACCTTGTATTTTCAGGAA 33
Db 87756 CTTGACCTTGTATTTTCAGGAA 87778
|||||

```

```

RESULT 11
AP002393/c 176107 bp DNA linear HTG 31-MAY-2000
LOCUS Homo sapiens chromosome 18 clone RP11-699C17 map 18q12, WORKING
DEFINITION DRAFT SEQUENCE, 19 unordered pieces.
ACCESSION AP002393
VERSION AP002393.1 GI:8131657
KEYWORDS HTG: HTGS PHASE1; HTGS DRAFT.
SOURCE Homo sapiens DNA, clone:RP11-699C17.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 176107)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Homo sapiens 176,107 genomic DNA of 18q12
Published Only in Database (2000)
2 (bases 1 to 176107)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (29-MAY-2000) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,
Japan (E-mail:hattori@gs.c.riken.go.jp,
URL:http://hqp.gsc.riken.go.jp/, Tel:81-42-778-9923,
Fax:81-42-778-9924)

```

```

REFERENCE
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
JOURNAL Published Only in Database (2000)
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (29-MAY-2000) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,
Japan (E-mail:hattori@gs.c.riken.go.jp,
URL:http://hqp.gsc.riken.go.jp/, Tel:81-42-778-9923,
Fax:81-42-778-9924)

```

## COMMENT

```

----- Genome Center
Center: RIKEN Genomic Sciences Center(GSC)
Center code: RIKEN
Web site: http://hqp.gsc.riken.go.jp/
Contact: hattori@gs.c.riken.go.jp
----- Project Information
Center project name: HumDraft18
Center Clone name: RP11-699C17
----- Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 163826 bases at least Q40
Consensus quality: 169603 bases at least Q30
Consensus quality: 172435 bases at least Q20
Insert size: 174307; sum-of-ctigs
Quality coverage: 5.63x in Q20 bases; sum-of-ctigs
-----
NOTE: This is a 'working draft' sequence. It currently consists of
19 contigs. The true order of the pieces is not known and their
order in this sequence record is arbitrary. Gaps between the
contigs are represented as runs N, but the exact sizes of the gaps
are unknown. This record will be updated with the finished sequence
as soon as it is available and the accession number will be
preserved
1 31231 contig of 31231 bp in length
31332 61792 contig of 30461 bp in length
61893 88918 contig of 27026 bp in length

```

```

89019 98085 contig of 9067 bp in length
98186 109413 contig of 11228 bp in length
109514 119559 contig of 10046 bp in length
119660 128623 contig of 8964 bp in length
128724 136240 contig of 7517 bp in length
136341 144191 contig of 7851 bp in length
144292 149366 contig of 5665 bp in length
150057 156749 contig of 6693 bp in length
156850 162315 contig of 5466 bp in length
162416 164672 contig of 2257 bp in length
164773 167491 contig of 2719 bp in length
167592 170792 contig of 3201 bp in length
170793 170892 contig of 100 bp
170893 172798 contig of 1906 bp in length
172799 172898 contig of 100 bp
172899 173274 contig of 376 bp in length
173275 173374 contig of 100 bp
173375 174909 contig of 1535 bp in length
174910 175009 contig of 100 bp
175010 176107 contig of 1098 bp in length.

FEATURES
      source
      1. 176107
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /chromosome="18"
      /map="18q12"
      /clone="RP11-699C17"
      1. 31231
      /note="assembly-fragment"
      31332. 61792
      /note="assembly-fragment"
      61893. 88918
      misc_feature

```

```

/note="assembly-fragment"
89019. 98085
/note="assembly-fragment"
98186. 109413
/note="assembly-fragment"
109514. 119559
/note="assembly-fragment"
119660. 128623
/note="assembly-fragment"
128724. 136240
/note="assembly-fragment"
136341. 144191
/note="assembly-fragment"
144292. 149366
/note="assembly-fragment"
150057. 156749
/note="assembly-fragment"
156850. 162315
/note="assembly-fragment"
162416. 164672
/note="assembly-fragment"
164773. 167491
/note="assembly-fragment"
167592. 170792
/note="assembly-fragment"
170793. 170892
/note="assembly-fragment"
170893. 172798
/note="assembly-fragment"
172799. 172898
/note="assembly-fragment"
172899. 173274
/note="assembly-fragment"
173275. 173374
/note="assembly-fragment"
173375. 174909
/note="assembly-fragment"
174910. 175009
/note="assembly-fragment"
175010. 176107
/note="assembly-fragment"

BASE COUNT 46994 a 37813 c 39266 g 50228 t 1806 others

ORIGIN
Query Match 13.1% Score 23; DB 2; Length 176107;
Best Local Similarity 100.0%; Pred. No. 0.027;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 CTTGAGCTTGTATTTTCAGGAA 33
Db 34317 CTTGAGCTTGTATTTTCAGGAA 34295

RESULT 12
AC090246/c 176612 bp DNA linear HTG 21-AUG-2002
LOCUS Homo sapiens chromosome 18 clone RP11-699C17 map 18, *** SEQUENCING
DEFINITION IN PROGRESS ***, 2 unordered pieces.
ACCESSION AC090246
VERSION AC090246.7 GI:22381545
KEYWORDS HTG; HTGS_PHASE1; HTGS_FULLTOP; HTGS_ACTIVERIN.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 176612)
AUTHORS Birren, B., Nusbaum, C. and Lander, E.
TITLE Homo sapiens chromosome 18, clone RP11-699C17
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 176612)
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S.,
Barra, N., Bastien, V., Boguslavsky, L., Boukhalter, B., Brown, A.,
Camara, J., Campopiano, A., Choepel, Y., Colangelo, M., Collins, S.,
Collamore, A., Cooke, P., Dearrellano, K., Dewar, K., Diaz, J. S.,
Dodge, S., Fato, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J.,
Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grande-Pierre, N.,
Hagos, B., Hartford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
Jones, C., Karalas, A., LaRoque, K., Lamazares, R., Landers, T.,
Lehoczky, J., Levine, R., Liu, G., Maclean, C., Macdonald, P.,
Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K.,
McPheeters, R., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V.,

```

Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunhkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnupack, R., Seaman, S., Severy, P., Sounguez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Travers, M., Travers, N., Triggillo, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

**TITLE**  
**JOURNAL**  
**REFERENCE**  
**AUTHORS**

Submitted (17-FEB-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
 3 (bases 1 to 176612)  
 Birren, B., Nussbaum, C., Lander, E., All, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collimore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Fato, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M., Meldrum, J., Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunhkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schnupack, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

**TITLE**  
**JOURNAL**  
**COMMENT**

Submitted (21-APR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Aug 21, 2002 this sequence version replaced g1:22123591.  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

**REFERENCE**  
**AUTHORS**

Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR  
 Web site: http://www.seq.wi.mit.edu  
 Contact: sequence\_submissions@genome.wi.mit.edu

----- Project Information -----  
 Center project name: 112328  
 Center clone name: 699\_C\_17

----- NOTE: This is a 'working draft' sequence. It currently consists of 2 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 131563: contig of 131563 bp in length  
 \* 131564 131663: gap of 100 bp  
 \* 131664 176612: contig of 44949 bp in length.

**FEATURES**  
**SOURCE**

1. 176612  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="18"  
 /map="18"  
 /clone="RP11-699C17"  
 /clone\_lib="RP11-11 Human Male BAC"  
 BASE COUNT 46327 a 37044 c 41351 g 51790 t 100 others  
 ORIGIN

Query Match 13.1% Score 23; DB 2; Length 176612;  
 Best Local Similarity 100.0%; Pred. No. 0.027;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 CTTGACCTTGTATTTTCAGAA 33  
 Db 28978 CTTGACCTTGTATTTTCAGAA 28956

RESULT 13  
 AC105227/C  
 LOCUS  
 DEFINITION  
 AC105227  
 VERSION  
 AC105227.6 GI:20258546  
 KEYWORDS  
 SOURCE  
 ORGANISM

185568 bp DNA linear PRI 29-APR-2002  
 Homo sapiens chromosome 18, clone RP11-1030E3, complete sequence.  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 1 (bases 1 to 185568)  
 Birren, B., Linton, L., Nussbaum, C. and Lander, E.  
 Homo sapiens chromosome 18, clone RP11-1030E3  
 Unpublished  
 2 (bases 1 to 185568)  
 Birren, B., Linton, L., Nussbaum, C., Lander, E., All, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhgalter, B., Brown, A., Camarata, J., Campolano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collimore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Fato, S., Ferreira, P., Fitzgibbon, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heatford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Larocque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., Maclean, C., Macdonald, P., Major, J., Margulis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McPherson, R., Meldrum, J., Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunhkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnupack, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Triggillo, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

**TITLE**  
**JOURNAL**  
**REFERENCE**  
**AUTHORS**

Submitted (26-DEC-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
 3 (bases 1 to 185568)  
 Birren, B., Linton, L., Nussbaum, C., Lander, E., All, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B., Brown, A., Camarata, J., Campolano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collimore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Fato, S., Ferreira, P., Fitzgibbon, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Larocque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Margulis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunhkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnupack, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Triggillo, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

**TITLE**  
**JOURNAL**

Submitted (21-APR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE  
AUTHORS

4 (bases 1 to 185568)  
 Birren,B., Linton,L., Nusbaum,C., Lande,E., All,A., Allen,N.,  
 Anderson,S., Barina,N., Bastien,V., Bloom,T., Boguslavsky,L.,  
 Boukhgalter,B., Brown,A., Camarata,J., Campoliano,A., Chang,J.,  
 Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,  
 Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,  
 Fero,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,  
 Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,  
 Hagos,B., Horton,L., Hulme,W., Illay,I., Johnson,R., Jones,C.,  
 Kamat,A., Karatas,A., Kells,C., Laroque,K., Lamazares,R.,  
 Landers,T., Lehoczký,J., Levine,R., Linblad-Troh,K., Liu,G.,  
 Maclean,C., MacDonald,P., Major,J., Margulis,N., Matthews,C.,  
 McCarthy,M., McKean,P., McKernan,K., Mejdum,J., Menus,L.,  
 Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,  
 Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,  
 Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,  
 Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,  
 Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupack,R.,  
 Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,  
 Strauss,N., Sudamanian,A., Talamas,J., Testaye,S., Theodore,J.,  
 Topham,K., Travers,M., Travis,N., Triglilio,J., Vassiliev,H.,  
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,  
 Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE  
JOURNAL

Submitted (29-APR-2002) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Apr 21, 2002 this sequence version replaced g1:20128372.  
 All repeats were identified using RepeatMasker:

## COMMENT

Smilt, A.F.A. & Green, P. (1996:1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR  
 Web site: <http://www-seq.wi.mit.edu>  
 Contact: [sequence.submissions@genome.wi.mit.edu](mailto:sequence.submissions@genome.wi.mit.edu)  
 ----- Project Information  
 Center project name: L23094  
 Center clone name: 1030\_E\_3  
 -----

## FEATURES

source  
 1. 185568  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="18"  
 /map="18"  
 /clone\_lib="RP11-1030E3"  
 /clone\_1lb="RPC1-11 Human Male BAC"  
 1042..1066  
 repeat\_region  
 /rpt\_family="AT\_rich"  
 1445..1486  
 repeat\_region  
 /rpt\_family="(TTTA)n"  
 complement(1541..1793)  
 /rpt\_family="AluSx"  
 complement(1995..2286)  
 /rpt\_family="L1R4"  
 complement(2373..2429)  
 /rpt\_family="L1R4"  
 complement(2616..2927)  
 /rpt\_family="AluY"  
 complement(2928..3238)  
 /rpt\_family="AluSx"  
 3557..3648  
 repeat\_region  
 /rpt\_family="L2"  
 complement(3649..4052)  
 /rpt\_family="MSFB"  
 4053..4283  
 repeat\_region  
 /rpt\_family="L2"  
 5396..5724  
 repeat\_region  
 /rpt\_family="MER2"  
 5964..6103  
 repeat\_region  
 /rpt\_family="L1ME3"  
 6118..7124  
 repeat\_region  
 /rpt\_family="L1PA12"

repeat\_region complement(7125..7415)  
 /rpt\_family="AluY"  
 repeat\_region 7416..7577  
 /rpt\_family="L1PA12"  
 repeat\_region 7578..8204  
 /rpt\_family="L1PA14"  
 repeat\_region 8215..8449  
 /rpt\_family="L1ME3"  
 repeat\_region 8671..9139  
 /rpt\_family="L1ME3"  
 repeat\_region 9140..9171  
 /rpt\_family="(TTCC)n"  
 9172..10748  
 /rpt\_family="SVA"  
 10749..12165  
 /rpt\_family="L1ME3"  
 12152..12299  
 /rpt\_family="L1ME3"  
 repeat\_region 12303..12481  
 /rpt\_family="L1MD1"  
 repeat\_region 12502..12604  
 /rpt\_family="L1ME3"  
 13063..13094  
 /rpt\_family="AT\_rich"  
 repeat\_region complement(13333..14104)  
 /rpt\_family="L1PA7"  
 repeat\_region complement(14164..14291)  
 /rpt\_family="MIR3"  
 complement(14292..14525)  
 /rpt\_family="L2"  
 14526..15051  
 /rpt\_family="MER1A"  
 repeat\_region complement(15052..15536)  
 /rpt\_family="L2"  
 repeat\_region complement(15535..15608)  
 /rpt\_family="MIR"  
 15988..16291  
 /rpt\_family="AluSx"  
 repeat\_region complement(16697..17281)  
 /rpt\_family="L1MC5"  
 17282..17344  
 /rpt\_family="L1MC5"  
 repeat\_region 17312..17675  
 /rpt\_family="MER61C"  
 repeat\_region complement(17709..17768)  
 /rpt\_family="L1MC4a"  
 repeat\_region complement(17790..18317)  
 /rpt\_family="MIR1G1"  
 complement(18408..18630)  
 /rpt\_family="L1MC5"  
 repeat\_region complement(18774..18891)  
 /rpt\_family="L1MC5"  
 repeat\_region complement(18899..19462)  
 /rpt\_family="L1R6A"  
 complement(19463..19751)  
 /rpt\_family="HERV871"  
 repeat\_region complement(19743..22357)  
 /rpt\_family="HERV871"  
 repeat\_region complement(22351..22457)  
 /rpt\_family="HERV871"  
 22458..23434  
 /rpt\_family="L1R5"

Query Match 13.1%; Score 23; DB 9; Length 185568;  
 Best Local Similarity 100.0%; Pred. No. 0.027;  
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 CTTGACCTTGTTGTTTCAGGAA 33  
 |||||  
 Db 164413 CTTGACCTTGTTGTTTCAGGAA 164391

RESULT 14



AC103349/c AC103349 167691 bp DNA linear HTG 21-AUG-2002  
 LOCUS Mus musculus clone RP23-303K3, WORKING DRAFT SEQUENCE, 19 unordered  
 DEFINITION pieces.  
 AC103349  
 AC103349.2 GI:22381452  
 VERSION HTG: HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULLLTOP.  
 KEYWORDS house mouse.  
 SOURCE Mus musculus  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus;  
 1 (bases 1 to 167691)  
 Birren, B., Nusbaum, C. and Lander, E.  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 2 (bases 1 to 167691)  
 Birren, B., Linton, L., Nusbaum, C., Lander, E., All, A., Allen, N.,  
 Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Bouknight, B.,  
 Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,  
 Choquel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A.,  
 Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S.,  
 Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,  
 Glade, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,  
 Hagos, B., Heaford, A., Horton, L., Huime, W., Iliev, I., Johnson, R.,  
 Jones, C., Kamat, A., Karatas, A., Kells, C., Lacroque, K.,  
 Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,  
 Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,  
 McCarthy, M., McEwan, P., McKernan, K., McPheters, R., Meltrin, J.,  
 Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C.,  
 Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D.,  
 Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,  
 Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,  
 Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnupack, R.,  
 Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,  
 Straus, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,  
 Topham, K., Travers, M., Travis, N., Trigglio, J., Vassiliev, H.,  
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,  
 Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 3 (bases 1 to 167691)  
 Submitted (24-NOV-2001) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA

Center project name: L17657  
 Center clone name: 303\_K-3  
 ----- Summary Statistics  
 Sequencing vector: Plasmid; n/a; 100% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.960731  
 Consensus quality: 161237 bases at least Q40  
 Consensus quality: 164563 bases at least Q30  
 Consensus quality: 165460 bases at least Q20  
 Insert size: 163000; agarose-fp  
 Insert size: 165891; sum-of-contigs  
 Quality coverage: 8.8 in Q20 bases; agarose-fp  
 Quality coverage: 8.6 in Q20 bases; sum-of-contigs  
 -----  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 19 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 \*  
 \* 1 4555: contig of 4555 bp in length  
 \* 4556 4655: gap of 100 bp  
 \* 4556 5264: contig of 609 bp in length  
 \* 5265 5364: gap of 100 bp  
 \* 5365 5944: contig of 580 bp in length  
 \* 5945 6044: gap of 100 bp  
 \* 6045 6444: contig of 1210 bp in length  
 \* 6445 7254: gap of 100 bp  
 \* 7255 7354: gap of 100 bp  
 \* 7355 8690: contig of 1336 bp in length  
 \* 8691 8790: gap of 100 bp  
 \* 8791 9904: contig of 1114 bp in length  
 \* 9905 10004: gap of 100 bp  
 \* 10005 11423: contig of 1417 bp in length  
 \* 11424 11521: gap of 100 bp  
 \* 11522 12995: contig of 1474 bp in length  
 \* 12996 13095: gap of 100 bp  
 \* 13096 14365: contig of 1270 bp in length  
 \* 14366 14465: gap of 100 bp  
 \* 14466 16284: contig of 1819 bp in length  
 \* 16285 16384: gap of 100 bp  
 \* 16385 18855: contig of 2471 bp in length  
 \* 18856 18955: gap of 100 bp  
 \* 18956 22999: contig of 4044 bp in length  
 \* 23000 23099: gap of 100 bp  
 \* 23100 26210: contig of 3111 bp in length  
 \* 26211 26310: gap of 100 bp  
 \* 26311 40871: contig of 14561 bp in length  
 \* 40872 40971: gap of 100 bp  
 \* 40972 56546: contig of 15575 bp in length  
 \* 56547 56646: gap of 100 bp  
 \* 56647 80881: contig of 24235 bp in length  
 \* 80882 80981: gap of 100 bp  
 \* 80982 103267: contig of 22286 bp in length  
 \* 103268 103367: gap of 100 bp  
 \* 103368 126955: contig of 23588 bp in length  
 \* 126956 127055: gap of 100 bp  
 \* 127056 167691: contig of 40636 bp in length.  
 Location/Qualifiers  
 1. 167691  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /clone="RP23-303K3"  
 /clone\_lib="RPCT-23 Female Mouse BAC"  
 1. 4555  
 /note="assembly\_fragment"  
 /clone\_end:SP6  
 vector\_side:left"  
 misc\_feature  
 misc\_feature  
 misc\_feature  
 misc\_feature  
 /note="assembly\_fragment"  
 5365. 5944  
 /note="assembly\_fragment"



\* 5563 5662: gap of unknown length  
\* 5663 6814: contig of 1152 bp in length  
\* 6815 6914: gap of unknown length  
\* 6915 8228: contig of 1314 bp in length  
\* 8229 9362: gap of unknown length  
\* 9363 9462: gap of unknown length  
\* 9463 11208: contig of 1746 bp in length  
\* 11209 11308: gap of unknown length  
\* 11309 12347: contig of 1039 bp in length  
\* 12348 12447: gap of unknown length  
\* 12448 13738: contig of 1291 bp in length  
\* 13739 13838: gap of unknown length  
\* 13839 14885: contig of 1047 bp in length  
\* 14886 14985: gap of unknown length  
\* 14986 16091: contig of 1106 bp in length  
\* 16092 16191: gap of unknown length  
\* 16192 17902: contig of 1711 bp in length  
\* 17903 18002: gap of unknown length  
\* 18003 19382: contig of 1380 bp in length  
\* 19383 19482: gap of unknown length  
\* 19483 20663: contig of 1181 bp in length  
\* 20664 20763: gap of unknown length  
\* 20764 21932: contig of 1169 bp in length  
\* 21933 22032: gap of unknown length  
\* 22033 23262: contig of 1230 bp in length  
\* 23263 25111: contig of 1749 bp in length  
\* 25112 25211: gap of unknown length  
\* 25212 26281: contig of 1070 bp in length  
\* 26282 26381: gap of unknown length  
\* 26382 27785: contig of 1404 bp in length  
\* 27786 27885: gap of unknown length  
\* 27886 29448: contig of 1563 bp in length  
\* 29449 29548: gap of unknown length  
\* 29549 31328: contig of 1780 bp in length  
\* 31329 31428: gap of unknown length  
\* 31429 33139: contig of 1711 bp in length  
\* 33140 33239: gap of unknown length  
\* 33240 34718: contig of 1479 bp in length  
\* 34719 34818: gap of unknown length  
\* 34819 36911: contig of 2093 bp in length  
\* 36912 37011: gap of unknown length  
\* 37012 38027: contig of 1016 bp in length  
\* 38028 38127: gap of unknown length  
\* 38128 39562: contig of 1435 bp in length  
\* 39563 39662: gap of unknown length  
\* 39663 40751: contig of 1089 bp in length  
\* 40752 40851: gap of unknown length  
\* 40852 42914: contig of 2063 bp in length  
\* 42915 43014: gap of unknown length  
\* 43015 44102: contig of 1088 bp in length  
\* 44103 44202: gap of unknown length  
\* 44203 45433: contig of 1231 bp in length  
\* 45434 45533: gap of unknown length  
\* 45534 46810: contig of 1277 bp in length  
\* 46811 46910: gap of unknown length  
\* 46911 48511: contig of 1601 bp in length  
\* 48512 48611: gap of unknown length  
\* 48612 50317: contig of 1706 bp in length  
\* 50318 50417: gap of unknown length  
\* 50418 52272: contig of 1855 bp in length  
\* 52273 52372: gap of unknown length  
\* 52373 53837: contig of 1465 bp in length  
\* 53838 53937: gap of unknown length  
\* 53938 55583: contig of 1646 bp in length  
\* 55584 55683: gap of unknown length  
\* 55684 57128: contig of 1445 bp in length  
\* 57129 57228: gap of unknown length  
\* 57229 58882: contig of 1654 bp in length  
\* 58883 58982: gap of unknown length  
\* 58983 61320: contig of 2338 bp in length  
\* 61321 61420: gap of unknown length

\* 61421 63901: contig of 2481 bp in length  
\* 63902 64001: gap of unknown length  
\* 64002 66344: contig of 2343 bp in length  
\* 66345 66444: gap of unknown length  
\* 66445 67984: contig of 1540 bp in length  
\* 67985 68084: gap of unknown length  
\* 68085 70337: contig of 2253 bp in length  
\* 70338 70437: gap of unknown length  
\* 70438 72788: contig of 2351 bp in length  
\* 72789 72888: gap of unknown length  
\* 72889 74331: contig of 1443 bp in length  
\* 74332 74431: gap of unknown length  
\* 74432 77106: contig of 2675 bp in length  
\* 77107 77206: gap of unknown length  
\* 77207 79143: contig of 1937 bp in length  
\* 79144 79243: gap of unknown length  
\* 79244 82097: contig of 2854 bp in length  
\* 82098 82197: gap of unknown length  
\* 82198 85184: contig of 2987 bp in length  
\* 85185 85284: gap of unknown length  
\* 85285 87051: contig of 1767 bp in length  
\* 87052 87151: gap of unknown length  
\* 87152 89984: contig of 2833 bp in length  
\* 89985 90084: gap of unknown length  
\* 90085 91616: contig of 1532 bp in length  
\* 91617 91716: gap of unknown length  
\* 91717 93274: contig of 1558 bp in length

## Query Match

Best Local Similarity 12.0%; Score 21; DB 2; Length 173702;  
100.0%; Pred. No. 0.43;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 GCAGGCCAGGCCCTTCTCTT 101

Db 147064 GCAGGCCAGGCCCTTCTCTT 147084

Search completed: July 28, 2003, 21:14:48  
Job time : 491.326 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 14:31:11 ; Search time 45.9617 Seconds  
(without alignments)  
8574.520 Million cell updates/sec

Title: US-10-010-408-1\_COPY\_1534\_1708  
Perfect score: 175  
Sequence: 1 AGTCCAGGAACTGAGCTTT.....GCTTAGATAAACACCCNAA 175

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapept 60.0

Searched: 2185239 seqs, 1125999159 residues  
Word size : 0

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

- Database : N.Geneseq\_101002.\*
- 1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*
  - 2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*
  - 3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*
  - 4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*
  - 5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*
  - 6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*
  - 7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*
  - 8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*
  - 9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*
  - 10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*
  - 11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*
  - 12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*
  - 13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*
  - 14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*
  - 15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*
  - 16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*
  - 17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*
  - 18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*
  - 19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*
  - 20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*
  - 21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*
  - 22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*
  - 23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*
  - 24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1          | 175   | 100.0       | 1708   | 20    | AAZ07516    |
| 2          | 121   | 69.1        | 439    | 24    | ABL59575    |
| 3          | 54    | 30.9        | 65     | 24    | ABN30189    |
| 4          | 35    | 20.0        | 1734   | 20    | AAZ76488    |
| 5          | 18    | 10.3        | 1668   | 24    | AAZ32413    |
| 6          | 18    | 10.3        | 1671   | 24    | ABL91207    |
| 7          | 18    | 10.3        | 1671   | 24    | ABL92605    |
| 8          | 18    | 10.3        | 2031   | 24    | ABL91770    |
| 9          | 18    | 10.3        | 2104   | 21    | AAF21412    |

|   |    |    |      |         |    |          |                      |
|---|----|----|------|---------|----|----------|----------------------|
| C | 10 | 18 | 10.3 | 2104    | 21 | AAA35290 | Human adenosine re   |
| C | 11 | 18 | 10.3 | 2104    | 22 | AAZ84222 | Human protein kina   |
| C | 12 | 18 | 10.3 | 2111    | 21 | AAZ75901 | DNA encoding a 60    |
| C | 13 | 18 | 10.3 | 2586    | 23 | AAZ56353 | Salmonella typhi D   |
| C | 14 | 18 | 10.3 | 5095    | 22 | AAH57492 | Human liver cell s   |
| C | 15 | 18 | 10.3 | 38644   | 21 | AAF21424 | Human low adenosin   |
| C | 16 | 18 | 10.3 | 38644   | 21 | AAZ35302 | Human adenosine re   |
| C | 17 | 18 | 10.3 | 1230025 | 20 | AAZ91990 | Nucleotide sequenc   |
| C | 18 | 17 | 9.7  | 24      | 20 | AAZ76528 | Mouse WISP-2 PCR p   |
| C | 19 | 17 | 9.7  | 317     | 21 | AAZ53400 | Arabidopsis thalia   |
| C | 20 | 17 | 9.7  | 387     | 22 | AAI80016 | Human polynucleoti   |
| C | 21 | 17 | 9.7  | 700     | 22 | AAH92347 | Human inflammatory   |
| C | 22 | 17 | 9.7  | 700     | 22 | AAH92348 | Human inflammatory   |
| C | 23 | 17 | 9.7  | 742     | 20 | AAZ61274 | Human signal pepti   |
| C | 24 | 17 | 9.7  | 791     | 22 | AAI96326 | Human neuroblastom   |
| C | 25 | 17 | 9.7  | 1026    | 24 | ABL42428 | Polypeptide-dihydr   |
| C | 26 | 17 | 9.7  | 1319    | 24 | ABO72690 | Human MDRT encodin   |
| C | 27 | 17 | 9.7  | 1485    | 23 | ABL05473 | Drosophila melanog   |
| C | 28 | 17 | 9.7  | 1521    | 21 | AAZ33310 | Arabidopsis thalia   |
| C | 29 | 17 | 9.7  | 2126    | 22 | AAZ46172 | Human DNA encoding   |
| C | 30 | 17 | 9.7  | 2261    | 20 | AAZ09837 | Human membrane spa   |
| C | 31 | 17 | 9.7  | 2998    | 21 | AAZ38856 | Human Jurkat cell    |
| C | 32 | 17 | 9.7  | 2998    | 21 | AAZ38864 | Human Jurkat cell    |
| C | 33 | 17 | 9.7  | 3775    | 23 | ABL05472 | Drosophila melanog   |
| C | 34 | 17 | 9.7  | 6643    | 22 | ABL6216  | Human nervous syst   |
| C | 35 | 17 | 9.7  | 10427   | 21 | AAZ36325 | Mechanical stress    |
| C | 36 | 17 | 9.7  | 10672   | 23 | ABL27119 | Drosophila melanog   |
| C | 37 | 17 | 9.7  | 17761   | 22 | AAK79415 | Human immune/haema   |
| C | 38 | 17 | 9.7  | 17761   | 23 | ABK42709 | Genomic sequence #   |
| C | 39 | 17 | 9.7  | 66743   | 23 | ABL27118 | Drosophila melanog   |
| C | 40 | 17 | 9.7  | 66804   | 24 | ABK87050 | Human transporter    |
| C | 41 | 17 | 9.7  | 349980  | 22 | AAZ86431 | Pyrococcus abyssi    |
| C | 42 | 16 | 9.1  | 51      | 22 | AAI76526 | Human silent SNP c   |
| C | 43 | 16 | 9.1  | 108     | 22 | AAI55188 | Probe #23874 used    |
| C | 44 | 16 | 9.1  | 219     | 22 | AAZ35423 | Human cardiovascular |
| C | 45 | 16 | 9.1  | 229     | 24 | ABL67914 | Ovary cancer relat   |

ALIGNMENTS

|          |  |
|----------|--|
| RESULT 1 |  |
| AAZ07516 | AAZ07516 standard; cDNA; 1708 BP.  |
| AAZ07516 |  |
| AC       | AAZ07516;  |
| XX       |  |
| DT       | 26-NOV-1999 (first entry)  |
| XX       |  |
| DE       | Rat HICP polypeptide encoding cDNA.  |
| XX       |  |
| KW       | Heparin-induced CGN-like protein; HICP; cell-associated activity; ss;  |
| KW       | cardiovascular disorder; aberrant cell proliferation; fibrotic disorder.   |
| OS       | Rattus sp.   |
| XX       |  |
| PN       | W09947556-A2.  |
| XX       |  |
| PD       | 23-SEP-1999.   |
| XX       |  |
| PF       | 18-MAR-1999; 99WO-US05999.   |
| XX       |  |
| PR       | 19-MAR-1998; 98US-0044273.   |
| XX       |  |
| PA       | (TUFT ) TUFTS COLLEGE.   |
| XX       |  |
| PI       | Castelliot JJ;   |
| XX       |  |
| DR       | WPI: 1999-562060/47.   |
| XX       |  |
| PT       | P-PSDB; AAZ27434.  |
| XX       |  |
| PT       | Nucleic acid sequences encoding rat heparin-induced CGN-like protein, used in methods to identify modulators or in diagnostic applications - |

XX Claim 2; Fig 1; 108bp; English.  
XX  
PS  
CC This cDNA encodes a rat heparin-induced CCN-like protein (HICP) protein.  
CC Agents that stimulate or inhibit HICP protein activity or expression.  
CC Antisense HICP nucleic acid molecules and HICP antibodies, can be used to  
CC modulate cell-associated activity. HICP modulators can be used to treat  
CC disorders characterized by aberrant HICP protein activity or expression.  
CC Probes capable of hybridizing to HICP mRNA or antibodies specific for  
CC HICP can be used to detect HICP activity in a biological sample. HICP  
CC can be used to treat disorders, such as a cardiovascular or fibrotic  
CC disorder, characterized by aberrant cell proliferation.  
XX  
SQ Sequence 1708 BP; 362 A; 486 C; 478 G; 382 T; 0 other;  
Query Match 100.0%; Score 175; DB 20; Length 1708;  
Best Local Similarity 100.0%; Pred. No. 3e-79; Indels 0; Gaps 0;  
Matches 175; Conservative 0; Mismatches 0;  
QY 1 AGTCCAGGAAGTCTGAGCTTGTATTTTCAGGAATGCACATCTTAAAGCACTGCAAAAC 60  
Db 1534 AGTCCAGGAAGTCTGAGCTTGTATTTTCAGGAATGCACATCTTAAAGCACTGCAAAAC 1593  
QY 61 AGAAGGCTCCACACCTCTGGCAGGCGCCCTTCTCTGTCAGCATGAGAAAGCAAG 120  
Db 1594 AGAAGGCTCCACACCTCTGGCAGGCGCCCTTCTCTGTCAGCATGAGAAAGCAAG 1653  
QY 121 GACAGAGAGTACTCTCTCTGGAGGACTAGTACCTGATTAACACCCAAA 175  
Db 1654 GACAGAGAGTACTCTCTCTGGAGGACTAGTACCTGATTAACACCCAAA 1708  
RESULT 2  
ABL59575  
ID ABL59575 standard; DNA; 439 BP.  
XX  
AC ABL59575;  
XX  
DT 17-JUL-2002 (first entry)  
XX  
DE Rat OST23 gene fragment SEQ ID NO:23.  
XX  
XX Rat; OST; osteoregenerative; parathyroid hormone; tibiae; osteopathic;  
KW gene therapy; parathyroid hormone receptor ligand; bone disorder;  
KW bone formation disorder; bone resorption disorder; osteoporosis;  
KW osteopenia; osteopetrosis; gene; ds.  
XX  
OS Rattus sp.  
XX  
XX WO200224943-A2.  
XX  
PD 28-MAR-2002.  
XX  
PF 19-SEP-2001; 2001MO-US29548.  
XX  
PR 19-SEP-2000; 2000US-233579P.  
XX  
PA (CURA-) CURAGEN CORP.  
PA (GLAX) GLAXO GROUP LTD.  
XX  
PI Horesovsky GJ, Noel LS, Raha D;  
XX  
DR WPI; 2002-401989/43.  
XX  
PT Identifying parathyroid hormone receptor ligands and osteoregenerative  
PT agents involves detecting the expression of nucleic acids which are  
PT regulated by parathyroid hormone  
XX  
PS Claim 45; Page 22; 90pp; English.  
XX  
CC The present invention describes a method (M1) for identifying parathyroid  
CC hormone receptor ligands (I) and osteoregenerative agents by contacting a  
CC test cell population (Cp) comprising cells expressing nucleic acid

CC sequences (S) of OST1-47 and 48, with a test agent, measuring nucleic  
CC acid sequence expression, comparing it with reference Cp and identifying  
CC the ligand and agent by the difference in expression levels. Also  
CC described is a method (M2) for treating a bone disorder in a subject by  
CC administering to the subject an agent that modulates the expression or  
CC activity of (S). (I) have osteopathic activities, and can be used to  
CC modulate the expression of OST 1-48, and can also be used in gene  
CC therapy. (M1) is useful for identifying parathyroid hormone receptor  
CC ligands (I) or osteoregenerative agents. OST gene sequence can be used  
CC for assessing the osteoregenerative activity of a test agent in a  
CC subject, and for diagnosing or determining the susceptibility to bone  
CC disorder and assessing efficacy of a treatment of a bone disorder in a  
CC subject, human or rodent. (M2) is useful for treating a bone disorder  
CC including bone formation disorder or bone resorption disorder (e.g.  
CC osteoporosis, osteopenia and osteopetrosis). OST polypeptides are useful  
CC as immunogens to raise anti-OST antibodies (II). Parathyroid hormone  
CC receptor ligands (I) identified by (M1) are useful for treating bone  
CC disorders. The present sequence represents an OST23 gene fragment  
CC isolated from rat tibiae, from the present invention.  
XX  
SQ Sequence 439 BP; 128 A; 108 C; 99 G; 104 T; 0 other;  
Query Match 69.1%; Score 121; DB 24; Length 439;  
Best Local Similarity 100.0%; Pred. No. 8.8e-52; Indels 0; Gaps 0;  
Matches 121; Conservative 0; Mismatches 0;  
QY 1 AGTCCAGGAAGTCTGAGCTTGTATTTTCAGGAATGCACATCTTAAAGCACTGCAAAAC 60  
Db 285 AGTCCAGGAAGTCTGAGCTTGTATTTTCAGGAATGCACATCTTAAAGCACTGCAAAAC 344  
QY 61 AGAAGGCTCCACACCTCTGGCAGGCGCCCTTCTCTGTCAGCATGAGAAAGCAAG 120  
Db 345 AGAAGGCTCCACACCTCTGGCAGGCGCCCTTCTCTGTCAGCATGAGAAAGCAAG 404  
QY 121 G 121  
Db 405 G 405  
RESULT 3  
ABN30189  
ID ABN30189 standard; DNA; 65 BP.  
XX  
AC ABN30189;  
XX  
DT 15-JUL-2002 (first entry)  
XX  
DE Rat spliced transcript detection oligonucleotide SEQ ID NO:2937.  
XX  
XX Human; mouse; rat; splice transcript; detection; RNA transcript;  
KW splice variant; transcriptome; oligonucleotide library; ss.  
XX  
OS Rattus norvegicus.  
XX  
XX WO200210449-A2.  
XX  
PD 07-FEB-2002.  
XX  
PF 20-JUL-2001; 2001MO-IB01903.  
XX  
PR 28-JUL-2000; 2000US-221607P.  
PR 02-MAY-2001; 2001US-287724P.  
XX  
PA (COMP-) COMPUGEN INC.  
XX  
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;  
XX  
DR WPI; 2002-257383/30.  
XX  
PT New oligonucleotide libraries comprising oligonucleotides which  
PT selectively hybridize to mRNAs transcribed from a transcription unit of  
PT a genome, useful for detecting tissue-, pathology-, and  
PT developmental-specific genes

```
XX XX Example 1; SEQ ID 2937; 47pp; English.
PS CC The present invention describes oligonucleotide libraries for detecting
XX CC messenger RNAs that populate a (sub-)transcriptome, where the
CC CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC CC transcription units that populate a genome. The library comprises
CC CC several oligonucleotides, each capable of hybridising selectively to a
CC CC set of messenger RNAs transcribed from a given transcription unit of
CC CC the genome, which encodes one or more messenger RNA splice variants.
CC CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC CC biological sample, in expression profiling studies, in qualitatively or
CC CC quantitatively characterising the corresponding transcriptome, and in
CC CC detecting RNA transcripts and splice variants of human or animal
CC CC transcriptomes. The libraries may also be used as specialised mini
CC CC libraries to detect transcripts of a sub-transcriptome under a
CC CC particular biological or pathological state, and so allowing the
CC CC detection of tissue- and pathology-specific genes such as those genes
CC CC only expressed in specific tissue under a specific pathological
CC CC condition; to detect developmental specific genes; and to detect RNA
CC CC transcripts and splice variants of a transcriptome of a patient suffering
CC CC from a particular disorder. ABN27253 to ABN59589 represent
CC CC oligonucleotide sequences from rats, humans and mice, which are used in
CC CC the exemplification of the present invention.
CC CC N.B. The sequence data for this patent did not form part of the printed
CC CC specification, but was obtained in electronic format directly from WIPO
CC CC at ftp.wipo.int/pub/published_pct_sequences.
XX XX
XX XX Sequence 65 BP; 21 A; 20 C; 13 G; 11 T; 0 other;
SQ
Query Match 30.9%; Score 54; DB 24; Length 65;
Best Local Similarity 100.0%; Pred. No. 1e-17;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 26 TTTCGAGATGCACTCTTTAAGCACTGCGCAAAAGAGAGCGCCACCTCT 79
Db 1 TTTCGAGATGCACTCTTTAAGCACTGCGCAAAAGAGAGCGCCACCTCT 54
RESULT 4
AAAX76488
ID AAX76488 standard; DNA; 1734 BP.
XX
XX AAX76488;
XX
XX 06-AUG-1999 (first entry)
XX
XX Mouse WISP-2 protein nucleotide sequence SEQ ID NO:17.
XX
XX WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;
XX KW connective tissue growth factor; cancer; melanoma; arteriosclerosis;
XX KW leukaemia; lymphoid malignancy; haematopoiesis-related disorder;
XX KW tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;
XX KW kidney disorder; bone-related disorder; osteoporosis; trauma; burn;
XX KW connective tissue disorder; cataplectic state; inflammation;
XX KW testicular-related disorder; angiogenesis; immunological disorder; ss.
XX
XX Mus sp.
XX
XX WO9921998-A1.
XX
XX 06-MAY-1999.
XX
XX 29-OCT-1998; 98WO-US22991.
XX
XX 14-APR-1998; 98US-0081695.
XX PR 29-OCT-1998; 97US-0063704.
XX PR 03-FEB-1998; 98US-0073612.
XX
XX (GETH ) GENENTECH INC.
XX
XX Bozstein DA, Cohen RL, Goddard A, Gurney AL, Hillan K;
XX Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WI;
PI
```

```
XX XX WPI: 1999-337420/28.
DR DR P-PSDB: AAX17651.
XX
XX New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3
XX
XX Example 2; Page 178-179; 284pp; English.
PS
XX
XX The present invention describes Wnt-1 induced secreted polypeptides,
CC CC WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2
CC CC and WISP-3 have homology to connective tissue growth factor (CTGF).
CC CC Products from the present invention can be used to treat WISP-related
CC CC disorders such as breast, ovarian, and colon cancer or melanoma. The
CC CC products can be used to treat arteriosclerosis. The products can also be
CC CC used to treat other diseases e.g. benign and malignant tumours,
CC CC leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,
CC CC hypothalamic and other glandular, macrophagal, epithelial, stromal, and
CC CC blastocoeleic disorders, haematopoiesis-related disorders, tissue-growth
CC CC disorders, skin disorders, desmoplasia, fibrotic lesions, kidney
CC CC disorders, bone-related disorders such as osteoporosis, trauma such as
CC CC burns, incisions, and other wounds, connective tissue disorders,
CC CC cataplectic states, testicular-related disorders, and inflammatory,
CC CC angiogenic and immunologic disorders including arteriosclerosis. The
CC CC products can also be used for detection and diagnosis especially of
CC CC individuals with neoplastic cell growth or proliferation. The products
CC CC can be used in the production of transgenic or knock-out animals.
CC CC Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing
CC CC cells.
XX
XX
XX Sequence 1734 BP; 355 A; 491 C; 495 G; 393 T; 0 other;
SQ
Query Match 20.0%; Score 35; DB 20; Length 1734;
Best Local Similarity 100.0%; Pred. No. 5.1e-08;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 77 TCTGGCAGGCCAGGGGCTTCTCTTCAGCATGAGA 111
Db 1611 TCTGGCAGGCCAGGGGCTTCTCTTCAGCATGAGA 1645
RESULT 5
AAD32413/C
ID AAD32413 standard; DNA; 1668 BP.
XX
XX AAD32413;
XX
XX 18-JUN-2002 (first entry)
XX
XX Chlamydia pneumoniae AR039 omcB/ompB gene.
XX
XX Chlamydiaceae family; chronic infection; persistent infection; PVK; nlpD;
XX KW Cpn0585; regulatory pathway; biosynthetic pathway; ompA; ompB; hsp60;
XX KW lipopolysaccharide; cardiovascular system; respiratory tract; therapy;
XX KW genital tract; reproductive system; atherosclerotic tissue; macrophage;
XX KW multiple sclerosis; conjunctiva; prophylaxis; antibacterial; gene; ds.
XX
XX Chlamydia pneumoniae.
XX
XX Key location/Qualifiers
XX CDS 1..1168
XX
XX /tag= a
XX /product= "Chlamydia pneumoniae AR039 omcB/ompB protein"
XX /note= "CDS does not include stop codon"
XX
XX WO200214516-A1.
XX
XX 21-FEB-2002.
XX
XX 17-AUG-2001; 2001WO-AU01021.
XX
XX 18-AUG-2000; 2000AU-0009540.
XX
XX
```

PA (UYOH-) UNIV QUEENSLAND TECHNOLOGY.  
 PA (MATH/) MATHEWS S A.  
 XX Timms P;  
 XX WPI: 2002-269197/31.  
 DR P-PSDB; AAE20302.  
 XX  
 PT Detecting Chlamydial organism in its persistent phase by detecting  
 PT expression change of range of genes belonging to their respective  
 PT biosynthetic pathways when expression is compared to that of organism  
 PT in lytic phase -  
 XX  
 PS Disclosure: Page 168-170; 196pp; English.  
 XX  
 CC The invention relates to composition and methods for detecting organisms  
 CC of the Chlamydiaceae family, including species of Chlamydia and  
 CC Chlamydia, in the persistent phase of their developmental cycle and  
 CC for the diagnosis of chronic or persistent infections caused by such  
 CC organisms. The composition is useful for modulating the expression of  
 CC gene such as *pyk*, *nlpD*, *Cpn0585*, a gene belonging to same regulatory/  
 CC biosynthetic pathway and *ompA*, *ompB*, *hsp60*, a gene involved in  
 CC lipopolysaccharide biosynthesis. It is also useful for modulating the  
 CC level and/or functional activity of an expression product of these  
 CC genes, where the gene is present in an epithelial cell (selected from  
 CC cardiovascular system, respiratory tract, genital tract, reproductive  
 CC system or conjunctiva), macrophage, or a cell associated with  
 CC atherosclerotic tissue or associated with multiple sclerosis brain  
 CC tissue. The composition is useful for treatment and/or prophylaxis of a  
 CC chronic infection caused by an organism of the Chlamydiaceae family in  
 CC a patient. Antigen associated with the persistent phase of the  
 CC developmental cycle of an organism of the Chlamydiaceae family, is  
 CC useful in the manufacture of a medicament, for treating and/or  
 CC preventing Chlamydiaceae infection in a patient. The present sequence  
 CC is Chlamydia pneumoniae AR039 *ompB/ompB* gene.  
 CC  
 XX  
 SO Sequence 1668 BP; 486 A; 342 C; 370 G; 470 T; 0 other;  
 XX  
 QY Query Match 10.3%; Score 18; DB 24; Length 1668;  
 Best Local Similarity 100.0%; Pred. NO. 23;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 XX  
 Db 3 TCCAGGAACTTGAGCTTT 20  
 1227 TCCAGGAACTTGAGCTTT 1210  
 XX  
 RESULT 6  
 ABL91207/C  
 ID ABL91207 standard; DNA; 1671 BP.  
 AC ABL91207;  
 XX  
 DT 29-JUL-2002 (first entry)  
 XX  
 DE Chlamydia pneumoniae cp6849 ORF DNA, SEQ ID NO:48.  
 XX  
 XX Chlamydial infection; antigen; immunogen; vaccine; diagnosis;  
 KW human respiratory disease; cardiovascular disease; atherosclerosis;  
 KW coronary artery disease; carotid artery stenosis; myocardial infarction;  
 KW cerebrovascular disease; aortic aneurysm; claudication; stroke;  
 KW strain CML029; open reading frame; ORF; gene; ds.  
 XX  
 OS Chlamydia pneumoniae.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..1671  
 FT /tag- a  
 FT /product- "cp6849"  
 FT sig\_peptide 1..66  
 FT /tag- b  
 FT mat\_peptide 67..1668  
 FT /tag- c

FT /product- "Mature protein"  
 XX  
 PN WO200202606-A2.  
 XX  
 XX 10-JAN-2002.  
 PD  
 XX  
 XX 03-JUL-2001; 2001WO-IB01445.  
 PE  
 XX 03-JUL-2000; 2000GB-0016363.  
 PR 11-JUL-2000; 2000GB-0017047.  
 PR 21-JUL-2000; 2000GB-0017983.  
 PR 07-AUG-2000; 2000GB-0019368.  
 PR 18-AUG-2000; 2000GB-0020440.  
 PR 14-SEP-2000; 2000GB-0022583.  
 PR 10-NOV-2000; 2000GB-0027549.  
 PR 22-DEC-2000; 2000GB-0031706.  
 XX  
 PA (CHIR-) CHIRON SPA.  
 XX  
 XX Ratl G, Grandl G;  
 PI  
 XX  
 DR WPI: 2002-154726/20.  
 DR N-PSDB; ABB90549.  
 XX  
 XX Novel Chlamydia pneumoniae protein useful in the manufacture of a  
 PT medicament for treatment or prevention of infection due to Chlamydia,  
 PT preferably Chlamydia pneumoniae, and for diagnostic purposes -  
 XX  
 PS Claim 5: Page 66; 364pp; English.  
 XX  
 XX Sequences ABB90526-ABB90715 represent novel proteins from Chlamydia  
 CC pneumoniae (strain CML029), and ABL91184-ABL91373 represent DNA encoding  
 CC them. The proteins are predicted to be immunogenic and may therefore be  
 CC useful in vaccine production and for diagnostic purposes. Chlamydia  
 CC pneumoniae is a common cause of respiratory disease in humans, and is  
 CC also involved in the development of cardiovascular diseases such as  
 CC atherosclerosis, coronary artery disease, carotid artery stenosis,  
 CC myocardial infarction, cerebrovascular disease, aortic aneurysm,  
 CC claudication and stroke. The proteins and nucleic acids of the invention  
 CC may be used in vaccines and pharmaceutical compositions for the  
 CC prevention or treatment of chlamydial infections, particularly Chlamydia  
 CC pneumoniae infections. The proteins may also be used in the detection of  
 CC Chlamydia pneumoniae, and the nucleic acids may be used in PCR, branched  
 CC DNA probe assay or blotting techniques for determining Chlamydia  
 CC pneumoniae gene expression. The present sequence represents a  
 CC specifically claimed DNA which encodes a Chlamydia pneumoniae protein of  
 CC the invention.  
 CC  
 XX  
 SO Sequence 1671 BP; 488 A; 342 C; 370 G; 471 T; 0 other;  
 XX  
 QY Query Match 10.3%; Score 18; DB 24; Length 1671;  
 Best Local Similarity 100.0%; Pred. NO. 23;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 XX  
 Db 3 TCCAGGAACTTGAGCTTT 20  
 1227 TCCAGGAACTTGAGCTTT 1210  
 XX  
 RESULT 7  
 ABL92605/C  
 ID ABL92605 standard; DNA; 1671 BP.  
 AC ABL92605;  
 XX  
 DT 05-JUN-2002 (first entry)  
 XX  
 DE Chlamydia pneumoniae DNA sequence SEQ ID NO:378.  
 XX  
 XX Chlamydial infection; Chlamydia; vaccine; detection; diagnosis;  
 KW antigen; antibacterial; immunostimulant; immune response;  
 KW Chlamydia-specific T-cell response; gene; ds.  
 XX



OS Chlamydia pneumoniae.  
XX  
PN WO200208267-A2.  
XX  
PD 31-JAN-2002.  
XX  
PF 20-JUL-2001; 2001WO-US23121.  
XX  
PR 20-JUL-2000; 2000US-0620412.  
PR 23-APR-2001; 2001US-0841132.  
XX  
PA (CORI-) CORIXA CORP.  
XX  
PI Fling SP, Skelky YAM, Probst P, Bhatia A;  
XX WPI; 2002-1179901/23.  
XX  
PT Novel compositions comprising Chlamydia Cap1 protein and its use in the  
XX treatment of Chlamydia infection -  
XX  
PS Disclosure; Page 333; 537pp; English.  
XX  
CC The present invention describes compositions comprising a Chlamydia Cap1  
CC protein and methods for the diagnosis and therapy of Chlamydia infection.  
CC Chlamydia DNA and protein sequences from the present invention can have  
CC antibacterial and immunostimulant activities, and can be used in  
CC vaccines. Compounds from the present invention can be used for eliciting  
CC an immune response, specifically stimulating a Chlamydia-specific T-cell  
CC response or inhibiting the development of a Chlamydia infection in an  
CC animal. Methods from the present invention can be used for detecting the  
CC presence of Chlamydia in a patient; to stimulate and/or expand T cells  
CC specific for a Chlamydia protein; and for treatment of a Chlamydia  
CC infection. ABL93394 to ABL92709 and ABB94096 to ABB94374 represent  
CC sequences used in the exemplification of the present invention.  
XX  
SQ Sequence 1671 BP; 488 A; 342 C; 370 G; 471 T; 0 other;  
XX  
Query Match 10.3%; Score 18; DB 24; Length 1671;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3 TCCAGGAACCTGACCTTT 20  
DB 1227 TCCAGGAACCTGACCTTT 1210  
XXXXXXXXXXXXXXXXXXXX  
RESULT 8  
ABL91770/c  
ID ABL91770 standard; DNA; 2031 BP.  
XX  
AC ABL91770;  
XX  
XX 28-MAY-2002 (first entry)  
XX  
DE Human polynucleotide SEQ ID NO 113.  
XX  
XX Human; HIV; HCV; gene expression; oligoribonucleotide; tumour; pathogen;  
KW Plasmodium; virus; viroid; cytokine; prion; antisense oligonucleotide;  
KW cytosolic; virucide; protozoacide; antibacterial; ds.  
XX  
OS Homo sapiens.  
XX  
XX DE10100586-C1.  
XX  
XX 11-APR-2002.  
XX  
XX 09-JAN-2001; 2001DE-1000586.  
XX  
XX 09-JAN-2001; 2001DE-1000586.  
XX  
XX (RIBO-) RIBOPHARMA AG.  
XX  
XX Kreutzer R, Limmer S, Rost S, Hadwiger P;  
PI

XX  
DR WPI; 2002-270454/32.  
XX  
PT Inhibiting gene expression in cells, useful for e.g. treating tumors,  
PT by introducing double-stranded complementary oligoRNA having unpaired  
PT terminal bases -  
XX  
XX Claim 13; Page 83-84; 104pp; German.  
XX  
CC The invention relates to a method for inhibiting expression of a target  
CC gene (ABL91658-ABL91797) in a cell by introducing at least one  
CC oligoribonucleotide that has a double-stranded structure consisting of at  
CC most 49 sequential nucleotide pairs, with at least part of one strand  
CC complementary with the target gene and has at least one end a  
CC single-stranded segment of 1-4 nt. The method provides  
CC oligoribonucleotides for antisense inhibition of gene expression useful  
CC e.g. for treating tumors but the oligoribonucleotides may also be  
CC directed against genes present in pathogens (e.g. Plasmodium or  
CC viruses/viroids), pathogenic on humans, animals or plants), or against  
CC cytokine, id, developmental or prion genes. The method provides more  
CC effective inhibition of gene expression than use of known  
CC oligonucleotides, probably because the unpaired overhang increases  
CC stability and thus intracellular concentration.  
XX  
SQ Sequence 2031 BP; 514 A; 549 C; 561 G; 407 T; 0 other;  
XX  
Query Match 10.3%; Score 18; DB 24; Length 2031;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 125 GCAGAGTACTCTCTCTG 142  
DB 1127 GCAGAGTACTCTCTCTG 1110  
XXXXXXXXXXXXXXXXXXXX  
RESULT 9  
AAF21412/c  
ID AAF21412 standard; DNA; 2104 BP.  
XX  
AC AAF21412;  
XX  
XX 14-MAR-2001 (first entry)  
XX  
DE Human low adenosine antisense oligonucleotide related sequence #2979.  
XX  
XX Low adenosine antisense oligonucleotide; phosphorothioate; allergy;  
KW human; airway disorder; bronchoconstriction; lung inflammation;  
KW surfactant depletion; respiratory; bronchodilator; antiinflammatory;  
KW immunosuppressive; antiallergic; hypotensive; cytosolic;  
KW respiratory obstruction; pulmonary obstruction; impeded respiration;  
KW surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;  
KW respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;  
KW pulmonary hypertension; emphysema; pulmonary transplantation rejection;  
KW chronic obstructive pulmonary disease; pulmonary infection; bronchitis;  
KW cancer; ss.  
XX  
OS Homo sapiens.  
XX  
XX WO200062736-A2.  
XX  
XX 26-OCT-2000.  
XX  
XX 24-MAR-2000; 2000WO-US08020.  
XX  
XX 06-APR-1999; 99US-0127958.  
XX  
XX (UYEC-) UNIV EAST CAROLINA.  
XX  
XX (NYCE/) NYCE J W.  
XX  
XX Nyce JW;  
XX  
XX WPI; 2000-679539/66.  
XX  
DR

PT Low adenosine (A) content antisense oligonucleotides which do not  
 PT trigger adenosine receptors during metabolism, useful e.g. for treating  
 PT cancers and respiratory obstructions  
 XX  
 PS Disclosure: Page 1393; 1592pp; English.  
 XX  
 CC The present invention describes low adenosine (A) content antisense  
 CC oligonucleotides and compositions (I) comprising them. In the antisense  
 CC oligonucleotides the A is replaced by a 'Universal' or alternative base.  
 CC (I) can have respiratory, bronchodilator, antiinflammatory, analgesic,  
 CC immunosuppressive, antiallergic, hypotensive and cytostatic activities.  
 CC The antisense oligonucleotides and (I) can be used to down-regulate the  
 CC expression and/or activity of target polypeptides associated with  
 CC lung/respiratory disorders and malignancies, such as stimulating and  
 CC activating peptide factors and transmitters, transcription factors,  
 CC immunoglobulins and antibodies, antibody receptors, cytokines and  
 CC chemokines, endogenously produced specific and non-specific enzymes,  
 CC binding proteins, adhesion molecules and their receptors, cytokine and  
 CC chemokine receptors, adenosine receptors, bradykinin receptors, central  
 CC nervous system (CNS) and peripheral nervous and non-nervous system  
 CC receptors, CNS and peripheral nervous and non-nervous system peptide  
 CC transmitters, defensins, growth factors, vasoreactive peptides and  
 CC receptors, binding proteins and malignancy associated proteins. The  
 CC antisense oligonucleotides may be used in this way to treat disorders  
 CC including respiratory obstruction (especially pulmonary obstruction  
 CC and/or bronchoconstriction) and/or lung inflammation, allergies)  
 CC and/or surfactant hypoproduction which are associated with a disease or  
 CC condition selected from pulmonary vasoconstriction, inflammation,  
 CC allergies, asthma, impaired respiration, respiratory distress syndrome  
 CC (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary  
 CC hypertension, emphysema, chronic obstructive pulmonary disease (COPD),  
 CC pulmonary transplantation rejection, pulmonary infections, bronchitis,  
 CC and/or cancer. AAF18434 to AAF21543 represent human polynucleotide  
 CC fragments and antisense oligonucleotides used in the exemplification of  
 CC the present invention.  
 XX  
 SO Sequence 2104 BP; 524 A; 581 C; 582 G; 417 T; 0 other;  
 Query Match 10.3%; Score 18; DB 21; Length 2104;  
 Best Local Similarity 100.0%; Pred. No. 23;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 125 GCAGAGTACTCTCTCTG 142  
 DB 1185 GCAGAGTACTCTCTCTG 1168  
 RESULT 10  
 AAA35290/C  
 ID AAA35290 standard; DNA: 2104 BP.  
 XX  
 AC AAA35290;  
 XX  
 DT 28-JUL-2000 (first entry)  
 XX  
 DE Human adenosine receptor related polynucleotide 2nd SEQ ID NO:164.  
 XX  
 KW Human; adenosine receptor; low adenosine antisense oligonucleotide;  
 KW phosphorothioate; impaired respiration; inflammation; allergy;  
 KW allergic disease; bronchoconstriction; inhibitor; antiinflammatory;  
 KW antiallergic; antiallergic; cytostatic; analgesic; impaired airway;  
 KW lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;  
 KW respiratory distress syndrome; pain; cystic fibrosis; emphysema;  
 KW pulmonary hypertension; chronic obstructive pulmonary disease; COPD;  
 KW cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200009525-A2.  
 XX  
 PD 24-FEB-2000.  
 XX  
 PF 03-AUG-1999; 99WO-US17712.

XX  
 PR 03-AUG-1999; 98US-0095212.  
 XX  
 PA (UYEC-) UNIV EAST CAROLINA.  
 XX  
 PI Nyce JW;  
 XX  
 DR WPI: 2000-205971/18.  
 XX  
 PT New antisense oligonucleotides useful for treating e.g. pulmonary  
 PT vasoconstriction, inflammation, allergies, asthma, hypertension,  
 PT bronchitis, emphysema, respiratory distress syndrome, ischemia or  
 PT cancers  
 XX  
 PS Disclosure: Page 1308; 1343pp; English.  
 XX  
 CC The present invention describes a new composition comprising an  
 CC antisense oligonucleotide (ON) with low adenosine (up to 15%), which  
 CC targets nucleic acids involved in bronchoconstriction, allergies, and/or  
 CC inflammation. The ON can have antiinflammatory, antiallergic,  
 CC antiallergic, cytostatic and analgesic activities. The compositions are  
 CC useful for the treatment of diseases associated with inflammation,  
 CC impaired airways, including lung disease and diseases whose secondary  
 CC effects afflict the lungs of a subject. They can be used for treating  
 CC e.g. ischaemic conditions, pulmonary vasoconstriction, allergies,  
 CC asthma, impaired respiration, respiratory distress syndrome, pain, cystic  
 CC fibrosis, pulmonary hypertension, emphysema, chronic obstructive  
 CC pulmonary disease (COPD), and cancers such as leukemias, lymphomas,  
 CC carcinomas, and cancers which may metastasise to the lungs, including  
 CC breast and prostate cancer. The reduction of the adenosine content of  
 CC the ONs reduces side effects. The A-containing ONs break down with the  
 CC release of deoxyadenosine which activates adenosine receptors causing the  
 CC bronchoconstriction and inflammation. AAA32313 to AAA35312 represent the  
 CC nucleotide sequences given in the sequence listing from the present  
 CC invention, which correspond to SEQ ID NO:1 to 2815, and then the last  
 CC 185 sequences are also called SEQ ID NO:1 to 185, but the sequences  
 CC differ from the previously named sequences. SEQ ID NO:11 to 1680  
 CC (AAA32323 to AAA33922) are specifically claimed ONs from the present  
 CC invention. N.B. Sequences given in the disclosure of the present  
 CC invention do not match up with their corresponding SEQ ID NO: sequences  
 CC given in the sequence listing.  
 XX  
 SO Sequence 2104 BP; 524 A; 581 C; 582 G; 417 T; 0 other;  
 Query Match 10.3%; Score 18; DB 21; Length 2104;  
 Best Local Similarity 100.0%; Pred. No. 23;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 125 GCAGAGTACTCTCTCTG 142  
 DB 1185 GCAGAGTACTCTCTCTG 1168  
 RESULT 11  
 AAC84222/C  
 ID AAC84222 standard; cDNA: 2104 BP.  
 XX  
 AC AAC84222;  
 XX  
 DT 19-MAR-2001 (first entry)  
 XX  
 DE Human protein kinase C-delta (PKC-delta) encoding cDNA.  
 XX  
 KW Antisense; human; protein kinase C-delta; PKC-delta; antiproliferative;  
 KW cytostatic; immunosuppressive; antidiabetic; neuroprotective; cancer;  
 KW antiinflammatory; antineumatic; antiallergic; hepatotropic; hepatitis;  
 KW inflammatory bowel disease; multiple sclerosis; pancreatitis; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 59..2089  
 FT /tag= a

```

FT XX /product= "PKC-delta"
PN XX WO200070091-A1.
XX XX 23-NOV-2000.
PD XX
XX XX 11-MAY-2000; 2000WO-US13170.
PF XX
XX XX 18-MAY-1999; 99US-0313930.
PR XX
XX XX (ISIS-) ISIS PHARM INC.
PA XX
XX XX Dean NM;
PI XX
XX XX WPI: 2001-025027/03.
DR XX P-PSDB: AAB48037.
XX XX
XX XX Antisense oligonucleotide for modulating human protein kinase C-delta
PT and treating diseases such as cancer, diabetes, comprises nucleotides
PT specifically hybridizable with nucleic acid encoding the protein kinase
XX
XX
XX Example 2; Page 70-74; 84pp: English.
PS XX
XX XX The invention provides antisense oligonucleotides that specifically
CC hybridize with human protein kinase C-delta (PKC-delta) polynucleotide.
CC The antisense oligonucleotides are useful for modulating the expression
CC of human PKC-delta in cells or tissues, especially adipose tissues which
CC express human PKC-delta, for inhibiting hyperproliferation of cells and
CC treating or preventing an abnormal proliferative condition such as
CC leukemia. They are also useful for reducing an inflammatory response of
CC human cells and treating an animal having a disease or condition
CC associated with tumor necrosis factor-alpha (TNF-alpha), especially
CC overexpression of TNF-alpha, such as inflammatory or autoimmune disease
CC or a condition which is infectious, diabetes, inflammatory bowel
CC disease, multiple sclerosis, pancreatitis, rheumatoid arthritis,
CC hepatitis, atopic dermatitis or allograft rejection and for reducing
CC the blood glucose level in a human. The present sequence represents the
CC cDNA encoding the human PKC-delta polypeptide.
XX XX
SQ Sequence 2104 BP; 524 A; 581 C; 582 G; 417 T; 0 other;

Query Match      10.3%; Score 18; DB 22; Length 2104;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 125 GCAGAGTACTCTCTCTG 142
   |||||||
DB 1185 GCAGAGTACTCTCTCTG 1168

RESULT 12
AAAT5901/C
ID AAAAT5901 standard; DNA; 2111 BP.
XX
XX AAAAT5901;
AC XX
XX
XX 22-JAN-2001 (first entry)
DT XX
XX
XX DNA encoding a 60 kDa cysteine-rich membrane protein.
DE XX
XX Cysteine-rich membrane protein; Chlamydia infection; bronchitis;
KM community acquired pneumonia; upper respiratory tract infection; vaccine;
KM sinusitis; ss.
XX
XX Chlamydia pneumoniae.
OS XX
XX
XX Key Location/Qualifiers
FH 139..1809
FT /tag= a
FT /product= "membrane-rich protein"
XX

```

```

PN WO200055326-A1.
XX
XX 21-SEP-2000.
PD XX
XX 09-MAR-2000; 2000WO-CA00240.
PF XX
XX 12-MAR-1999; 99US-0123966.
PR XX
XX (AVET ) AVENTIS PASTEUR LTD.
XX
XX Murdin AD, Oomen RP, Wang J, Dunn P;
PI
XX WPI: 2000-618918/59.
DR
XX P-PSDB: AAB18820.
XX
XX New polynucleotides encoding a 60kDa cysteine-rich membrane protein
PT from Chlamydia, useful as a vaccine for preventing and treating
PT Chlamydia infection in mammals
XX
XX Claim 2; Fig 1; 77pp: English.
PS
XX
XX The present sequence encodes a Chlamydia 60 kDa cysteine-rich membrane
CC protein. The membrane-rich polynucleotide and polypeptide are useful
CC for preventing or treating Chlamydia infection, such as community
CC pneumonia, upper respiratory tract infections, bronchitis and sinusitis.
CC They are also useful for diagnosing Chlamydia infection by assaying a
CC body fluid of a mammal. The polypeptide is useful for vaccine
CC production.
XX
XX
SQ Sequence 2111 BP; 634 A; 417 C; 442 G; 618 T; 0 other;

Query Match      10.3%; Score 18; DB 21; Length 2111;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 TCCAGGAACCTGAGCTTT 20
   |||||||
DB 1365 TCCAGGAACCTGAGCTTT 1348

RESULT 13
AAS56353
ID AAS56353 standard; DNA; 2586 BP.
XX
XX AAS56353;
AC XX
XX
XX 13-FEB-2002 (first entry)
DT XX
XX
XX Salmonella typhi DNA for cellular proliferation protein #386.
DE
XX
XX Antisense; ds; prokaryotic cellular proliferation gene;
KM antibiotic; antibacterial; drug design.
KM
XX Salmonella typhi.
OS
XX
XX WO200170955-A2.
PN
XX
XX 27-SEP-2001.
PD
XX
XX 21-MAR-2001; 2001WO-US09180.
PF
XX 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
XX (ELIT-) ELITRA PHARM INC.
PA
XX
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;

```

XX WPI; 2001-611495/70.  
DR P-PSDB; A0038494.  
XX  
XX New polynucleotides for the identification and development of  
PT antibiotics, comprise sequences of antisense nucleic acids -  
XX  
PS Claim 27; Seq ID No 9990; 511pp; English.  
XX  
XX The invention relates to antisense inhibitors of genes essential to  
CC prokaryotic cellular proliferation, their use in identifying the  
CC genes, their use in the discovery of novel antibiotics, the essential  
CC genes themselves and the encoded proteins. The prokaryotes used are  
CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*  
CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The  
CC invention is also useful for the identification of potential new targets  
CC for antibiotic development. The antisense nucleic acids can also be used  
CC to identify proteins used in proliferation, to express these proteins,  
CC and to obtain antibodies capable of binding to the expressed proteins.  
CC The proteins can be used to screen compounds in rational drug discovery  
CC programmes. The antisense nucleic acid sequence is also useful to screen  
CC for homologous nucleic acids which are required for cell proliferation in  
CC a wide variety of organisms. The present sequence encodes an  
CC essential prokaryotic cellular proliferation protein.  
CC Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic  
CC format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 2586 BP; 642 A; 618 C; 764 G; 561 T; 1 other;  
  
Query Match 10.3%; Score 18; DB 23; Length 2586;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
OY 51 CTCGCAAAACAGGAGGC 68  
XXXXXXXXXXXXXXXXXXXX  
DB 808 CTCGCAAAACAGGAGGC 825  
  
RESULT 14  
AAH57492  
ID AAH57492 standard; cDNA; 5095 BP.  
XX  
AC AAH57492;  
XX  
DT 10-SEP-2001 (first entry)  
XX  
DE Human liver cell specific cDNA sequence SEQ ID NO:332.  
XX  
KW Human; tissue specific; diagnosis; brain; heart; skeletal muscle;  
KW lung; liver; uterus; ovary; stomach; intestine; kidney; pancreas; ss;  
KW metabolic disease; developmental disease; cytostatic; immunomodulatory;  
KW neuroprotective; gene therapy; cancer; immunopathology; neuropathology.  
XX  
OS Homo sapiens.  
XX  
PN WO200132927-A2.  
XX  
PD 10-MAY-2001.  
XX  
PF 02-NOV-2000; 2000MO-US30396.  
XX  
PR 04-NOV-1999; 99US-0163508.  
XX  
PA (INCY-) INCYTE GENOMICS INC.  
XX  
PI Sorinasse T, Selhammer JT, Watson GA;  
XX  
DR WPI; 2001-291057/30.  
XX  
PT New cell and tissue specific polynucleotides useful for diagnosis,  
PT prognosis or monitoring of treatments for disorders where the gene is

PT associated with a cancer, immunopathology or neuropathology -  
XX  
XX Claim 1; Page 251-252; 327pp; English.  
PS  
XX  
XX AAH57161 to AAH57576 represent cell and tissue specific polynucleotide  
CC sequences (I). (I) can have cytostatic, immunomodulatory and  
CC neuroprotective activities, and can be used in gene therapy. (I) and  
CC proteins (II) encoded by them are used in high throughput screening  
CC assays to select DNA molecules, RNA molecules, peptide nucleic acids,  
CC mimetics, peptides, proteins, agonists, antagonists, antibodies or  
CC their fragments, immunoglobulins, inhibitors, drug compounds and  
CC pharmaceutical agents. Expression of (I) in a sample indicates the  
CC differentiation of embryonic stem cells into a tissue selected from  
CC brain, heart, kidney, liver, lung, skeletal muscle or pancreatic  
CC tissues. (I) and (II) are used to produce an expression profile that  
CC defines a metabolic or developmental process, treatment, condition,  
CC disease or disorder. The gene profile can be used for diagnosis,  
CC prognosis or monitoring of treatments and for investigating a  
CC predisposition to a disorder where the gene is associated with a  
CC cancer, immunopathology or neuropathology.  
XX  
SQ Sequence 5095 BP; 1596 A; 867 C; 941 G; 1691 T; 0 other;  
  
Query Match 10.3%; Score 18; DB 22; Length 5095;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
OY 74 ACCTCGCAGGCGCAGG 91  
XXXXXXXXXXXXXXXXXXXX  
DB 3636 ACCTCGCAGGCGCAGG 3653  
  
RESULT 15  
AAE21424/c  
ID AAE21424 standard; DNA; 38644 BP.  
XX  
AC AAE21424;  
XX  
DT 14-MAR-2001 (first entry)  
XX  
DE Human low adenosine antisense oligonucleotide related sequence #2991.  
XX  
KW Low adenosine antisense oligonucleotide; phosphorothioate; allergy;  
KW human; airway disorder; bronchoconstriction; lung inflammation;  
KW surfactant depletion; respiratory; bronchodilator; antiinflammation;  
KW immunosuppressive; antiallergic; hypotensive; cytostatic;  
KW respiratory obstruction; pulmonary obstruction; impeded respiration;  
KW surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;  
KW respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;  
KW pulmonary hypertension; emphysema; pulmonary transplantation rejection;  
KW chronic obstructive pulmonary disease; pulmonary infection; bronchitis;  
KW cancer; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200062736-A2.  
XX  
PD 26-OCT-2000.  
XX  
PF 24-MAR-2000; 2000MO-US08020.  
XX  
PR 06-APR-1999; 99US-0127958.  
XX  
PA (UYEC-) UNIV EAST CAROLINA.  
XX  
PA (NYCE/) NYCE J W.  
XX  
PI Nyce JW;  
XX  
DR WPI; 2000-679539/66.  
XX  
PT Low adenosine (A) content antisense oligonucleotides which do not  
PT trigger adenosine receptors during metabolism, useful e.g. for treating  
PT cancers and respiratory obstructions -

XX  
PS Disclosure: Page 1400-1409; 1592pp; English.

CC The present invention describes low adenosine (A) content antisense  
CC oligonucleotides and compositions (I) comprising them. In the antisense  
CC oligonucleotides the A is replaced by a 'Universal' or alternative base.  
CC (I) can have respiratory, bronchodilator, antiinflammatory, analgesic,  
CC immunosuppressive, antiasthmatic, hypotensive and cytostatic activities.  
CC The antisense oligonucleotides and (I) can be used to down-regulate the  
CC expression and or activity of target polypeptides associated with  
CC lung/respiratory disorders and malignancies, such as stimulating and  
CC activating peptide factors and transmitters, transcription factors,  
CC immunoglobulins and antibodies, antibody receptors, cytokines and  
CC chemokines, endogenously produced specific and non-specific enzymes,  
CC binding proteins, adhesion molecules and their receptors, cytokine and  
CC chemokine receptors, adenosine receptors, bradykinin receptors, central  
CC nervous system (CNS) and peripheral nervous and non-nervous system  
CC receptors, CNS and peripheral nervous and non-nervous system peptide  
CC transmitters, defensins, growth factors, vasoactive peptides and  
CC receptors, binding proteins and malignancy associated proteins. The  
CC antisense oligonucleotides may be used in this way to treat disorders  
CC including respiratory obstruction (especially pulmonary obstruction  
CC and/or bronchoconstriction) and/or lung inflammation, allergy(ies)  
CC and/or surfactant hypoproduction which are associated with a disease or  
CC condition selected from pulmonary vasoconstriction, inflammation,  
CC allergies, asthma, impeded respiration, respiratory distress syndrome  
CC (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary  
CC hypertension, emphysema, chronic obstructive pulmonary disease (COPD),  
CC pulmonary transplantation rejection, pulmonary infections, bronchitis,  
CC and/or cancer. AAF18434 to AAF21543 represent human polynucleotide  
CC fragments and antisense oligonucleotides used in the exemplification of  
CC the present invention.

XX  
SQ Sequence 38644 BP; 9551 A; 10265 C; 10290 G; 8538 T; 0 other;

Query Match 10.3%; Score 18; DB 21; Length 38644;

Best Local Similarity 100.0%; Pred. No. 25;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 125 GCAGAGTACTCTCTCTG 142

Db 16814 GCAGAGTACTCTCTCTG 16797

Search completed: July 28, 2003, 15:54:59  
Job time : 47.9617 secs





APPLICANT: Hillan, Kenneth J.  
APPLICANT: Lawrence, David A.  
APPLICANT: Levine, Arnold J.  
APPLICANT: Pennica, Diane  
APPLICANT: Roy, Margaret Ann  
TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME  
FILE REFERENCE: P11762  
CURRENT APPLICATION NUMBER: US/09/182,145B  
CURRENT FILING DATE: 1998-10-29  
EARLIER APPLICATION NUMBER: US 60/063,704  
EARLIER FILING DATE: 1997-10-29  
EARLIER APPLICATION NUMBER: US 60/073,612  
EARLIER FILING DATE: 1998-02-04  
EARLIER APPLICATION NUMBER: US 60/081,695  
EARLIER FILING DATE: 1998-04-14  
NUMBER OF SEQ ID NOS: 156  
SEQ ID NO 18  
LENGTH: 1734  
TYPE: DNA  
ORGANISM: Mus musculus  
US-09-182-145-18

Query Match 20.0%; Score 35; DB 4; Length 1734;  
Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 TCTGCAGGCGCGCTTCTCTTCAGCATGAGA 114  
124 TCTGCAGGCGCGCTTCTCTTCAGCATGAGA 90

RESULT 3  
US-09-313-930-1/c  
Sequence 1, Application US/09313930  
Patent No. 6235723  
GENERAL INFORMATION:  
APPLICANT: Dean, Nicholas M.  
TITLE OF INVENTION: Antisense Oligonucleotide Modulation of Human Protein  
FILE REFERENCE: ISPH-0357  
CURRENT APPLICATION NUMBER: US/09/313,930  
CURRENT FILING DATE: 1999-05-18  
NUMBER OF SEQ ID NOS: 20  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 1  
LENGTH: 2104  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (59)..(2089)  
PUBLICATION INFORMATION:  
AUTHORS: Aris, J. P.  
AUTHORS: Basta, P. V.  
AUTHORS: Holmes, W. D.  
AUTHORS: Ballas, L. M.  
AUTHORS: Moomaw, C.  
AUTHORS: Rankl, N. B.  
AUTHORS: Blobel, G.  
AUTHORS: Loomis, C. R.  
AUTHORS: Burns, D. J.  
TITLE: Molecular and biochemical characterization of a  
TITLE: recombinant human PKC-delta family member  
JOURNAL: Biochim. Biophys. Acta  
VOLUME: 1174  
ISSUE: 2  
PAGES: 171-181  
DATE: 1993-08-19  
DATABASE ACCESSION NUMBER: L07860  
DATABASE ENTRY DATE: 1993-11-02  
US-09-313-930-1

Query Match 10.3%; Score 18; DB 4; Length 2104;  
Best Local Similarity 100.0%; Pred. No. 2;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 125 GCAGAGTACTCTCTCTG 142  
DB 1185 GCAGAGTACTCTCTCTG 1168

RESULT 4  
US-09-182-145-110/c  
Sequence 110, Application US/09182145B  
Patent No. 6387657  
GENERAL INFORMATION:  
APPLICANT: Botstein, David A.  
APPLICANT: Cohen, Robert  
APPLICANT: Goddard, Audrey  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: Lawrence, David A.  
APPLICANT: Levine, Arnold J.  
APPLICANT: Pennica, Diane  
APPLICANT: Roy, Margaret Ann  
TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME  
FILE REFERENCE: P11762  
CURRENT APPLICATION NUMBER: US/09/182,145B  
CURRENT FILING DATE: 1998-10-29  
EARLIER APPLICATION NUMBER: US 60/063,704  
EARLIER FILING DATE: 1997-10-29  
EARLIER APPLICATION NUMBER: US 60/073,612  
EARLIER FILING DATE: 1998-02-04  
EARLIER APPLICATION NUMBER: US 60/081,695  
EARLIER FILING DATE: 1998-04-14  
NUMBER OF SEQ ID NOS: 156  
SEQ ID NO 110  
LENGTH: 24  
TYPE: DNA  
ORGANISM: Artificial sequence  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 1-24  
OTHER INFORMATION: Sequence is synthesized.  
Patent No. 6387657  
US-09-182-145-110

Query Match 9.7%; Score 17; DB 4; Length 24;  
Best Local Similarity 100.0%; Pred. No. 5.1;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 TCTGCAGGCGCGCGCC 93  
DB 17 TCTGCAGGCGCGCGCC 1

RESULT 5  
US-08-966-316-2/c  
Sequence 2, Application US/08966316  
Patent No. 5932445  
GENERAL INFORMATION:  
APPLICANT: Lai, Preeti  
APPLICANT: Au-Yang, Janice  
APPLICANT: Reddy, Roopa  
APPLICANT: Murty, Lynn E.  
APPLICANT: Mathur, Preete  
TITLE OF INVENTION: SIGNAL PEPTIDE - CONTAINING PROTEINS  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA



ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/966.316  
FILING DATE: Herewith  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0424 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-855-0555  
TELEFAX: 650-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 742 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: COLNET02  
CLONE: 1457779  
US-08-966-316-2

Query Match 9.7%; Score 17; DB 2; Length 742;  
Best Local Similarity 100.0%; Pred. No. 6.3;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 78 CTGGCAGCGCAGGGCCT 94  
DB 428 CTGGCAGCGCAGGGCCT 412

RESULT 6  
US-09-556-877-55/c  
Sequence 55, Application US/09556877  
Patent No. 6432916  
GENERAL INFORMATION:  
APPLICANT: Probst, Peter  
APPLICANT: Bhatia, Ajay  
APPLICANT: Skelky, Yasir  
APPLICANT: Fling, Steve  
APPLICANT: Maisonneuve, Jeff  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND  
TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION  
FILE REFERENCE: 210121.469C5  
CURRENT APPLICATION NUMBER: US/09/556.877  
CURRENT FILING DATE: 2000-04-19  
NUMBER OF SEQ ID NOS: 305  
SOFTWARE: FASTSEQ for Windows Version 3.0/4.0  
SEQ ID NO 55  
LENGTH: 463  
TYPE: DNA  
ORGANISM: Chlamydia trachomatis  
US-09-556-877-55

Query Match 9.1%; Score 16; DB 4; Length 463;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 117 AAGGACAGCAGAGTA 132  
DB 412 AAGGACAGCAGAGTA 397

RESULT 7  
US-09-556-877-58/c  
Sequence 58, Application US/09556877  
Patent No. 6432916  
GENERAL INFORMATION:  
APPLICANT: Probst, Peter  
APPLICANT: Bhatia, Ajay  
APPLICANT: Skelky, Yasir  
APPLICANT: Fling, Steve  
APPLICANT: Maisonneuve, Jeff  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND  
TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION  
FILE REFERENCE: 210121.469C5  
CURRENT APPLICATION NUMBER: US/09/556.877  
CURRENT FILING DATE: 2000-04-19  
NUMBER OF SEQ ID NOS: 305  
SOFTWARE: FASTSEQ for Windows Version 3.0/4.0  
SEQ ID NO 58  
LENGTH: 463  
TYPE: DNA  
ORGANISM: Chlamydia trachomatis  
US-09-556-877-58

OY 117 AAGGACAGCAGAGTA 132  
DB 412 AAGGACAGCAGAGTA 397

RESULT 8  
US-09-620-412C-55/c  
Sequence 55, Application US/09620412C  
Patent No. 6448234  
GENERAL INFORMATION:  
APPLICANT: Steven P. Fling  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND  
TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION  
FILE REFERENCE: 210121.469C7  
CURRENT APPLICATION NUMBER: US/09/620.412C  
CURRENT FILING DATE: 2000-07-20  
NUMBER OF SEQ ID NOS: 363  
SOFTWARE: FASTSEQ for Windows Version 3.0/4.0  
SEQ ID NO 55  
LENGTH: 463  
TYPE: DNA  
ORGANISM: Chlamydia trachomatis  
US-09-620-412C-55

Query Match 9.1%; Score 16; DB 4; Length 463;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 117 AAGGACAGCAGAGTA 132  
DB 412 AAGGACAGCAGAGTA 397

RESULT 9  
US-09-620-412C-58/c  
Sequence 58, Application US/09620412C  
Patent No. 6448234  
GENERAL INFORMATION:  
APPLICANT: Steven P. Fling  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND  
TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION  
FILE REFERENCE: 210121.469C7  
CURRENT APPLICATION NUMBER: US/09/620.412C  
CURRENT FILING DATE: 2000-07-20  
NUMBER OF SEQ ID NOS: 363  
SOFTWARE: FASTSEQ for Windows Version 3.0/4.0

SEQ ID NO 58  
LENGTH: 463  
TYPE: DNA  
ORGANISM: Chlamydia trachomatis  
US-09-620-412C-58

Query Match 9.1%; Score 16; DB 4; Length 463;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 117 AAGGACAGCAGAGTA 132  
Db 412 AAGGACAGCAGAGTA 397

RESULT 10  
US-09-556-877-22/C

Sequence 22, Application US/09556877  
Patent No. 6432916  
GENERAL INFORMATION:  
APPLICANT: Probst, Peter  
APPLICANT: Bhatia, Ajay  
APPLICANT: Skelky, Yasir  
APPLICANT: Filing, Steve  
APPLICANT: Maisonneuve, Jeff  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND  
FILE REFERENCE: 210121.469C5  
CURRENT APPLICATION NUMBER: US/09/556,877  
NUMBER OF SEQ ID NOS: 305  
SOFTWARE: FastSeq for Windows Version 3.0/4.0  
SEQ ID NO 22  
LENGTH: 601  
TYPE: DNA  
ORGANISM: Chlamydia trachomatis  
US-09-556-877-22

Query Match 9.1%; Score 16; DB 4; Length 601;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 117 AAGGACAGCAGAGTA 132  
Db 537 AAGGACAGCAGAGTA 522

RESULT 11

US-09-288-594A-22/C  
Sequence 22, Application US/09288594A  
Patent No. 6447779  
GENERAL INFORMATION:  
APPLICANT: Probst, Peter  
APPLICANT: Bhatia, Ajay  
APPLICANT: Skelky, Yasir A.W.  
APPLICANT: Filing, Steven P.  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE TREATMENT  
FILE REFERENCE: 210121.469C1  
CURRENT APPLICATION NUMBER: US/09/288,594A  
NUMBER OF SEQ ID NOS: 49  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 22  
LENGTH: 601  
TYPE: DNA  
ORGANISM: Chlamydia trachomatis  
US-09-288-594A-22

Query Match 9.1%; Score 16; DB 4; Length 601;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 117 AAGGACAGCAGAGTA 132  
Db 537 AAGGACAGCAGAGTA 522

RESULT 12

US-09-620-412C-22/C  
Sequence 22, Application US/09620412C  
Patent No. 6448234  
GENERAL INFORMATION:  
APPLICANT: Steven P. Filing  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND  
FILE REFERENCE: 210121.469C7  
CURRENT APPLICATION NUMBER: US/09/620,412C  
CURRENT FILING DATE: 2000-07-20  
NUMBER OF SEQ ID NOS: 363  
SOFTWARE: FastSeq for Windows Version 3.0/4.0  
SEQ ID NO 22  
LENGTH: 601  
TYPE: DNA  
ORGANISM: Chlamydia trachomatis  
US-09-620-412C-22

Query Match 9.1%; Score 16; DB 4; Length 601;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 117 AAGGACAGCAGAGTA 132  
Db 537 AAGGACAGCAGAGTA 522

RESULT 13

US-09-556-877-25/C  
Sequence 25, Application US/09556877  
Patent No. 6432916  
GENERAL INFORMATION:  
APPLICANT: Probst, Peter  
APPLICANT: Bhatia, Ajay  
APPLICANT: Skelky, Yasir  
APPLICANT: Filing, Steve  
APPLICANT: Maisonneuve, Jeff  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND  
FILE REFERENCE: 210121.469C5  
CURRENT APPLICATION NUMBER: US/09/556,877  
CURRENT FILING DATE: 2000-04-19  
NUMBER OF SEQ ID NOS: 305  
SOFTWARE: FastSeq for Windows Version 3.0/4.0  
SEQ ID NO 25  
LENGTH: 696  
TYPE: DNA  
ORGANISM: Chlamydia trachomatis  
US-09-556-877-25

Query Match 9.1%; Score 16; DB 4; Length 696;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 117 AAGGACAGCAGAGTA 132  
Db 538 AAGGACAGCAGAGTA 523

RESULT 14

US-09-288-594A-25/C  
Sequence 25, Application US/09288594A  
Patent No. 6447779  
GENERAL INFORMATION:  
APPLICANT: Probst, Peter  
APPLICANT: Bhatia, Ajay  
APPLICANT: Skelky, Yasir A.W.

```

: APPLICANT: Fling, Steven P.
: TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE TREATMENT
: TITLE OF INVENTION: AND DIAGNOSIS OF CHLAMYDIAL INFECTION
: FILE REFERENCE: 210121.469C1
: CURRENT APPLICATION NUMBER: US/09/288,594A
: CURRENT FILING DATE: 1999-04-08
: NUMBER OF SEQ ID NOS: 49
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 25
: LENGTH: 696
: TYPE: DNA
: ORGANISM: Chlamydia trachomatis
US-09-288-594A-25

```

```

Query Match          9.1%; Score 16; DB 4; Length 696;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 117 AAGGACACGACAGTA 132
    |||||||
Db 538 AAGGACACGACAGTA 523

```

```

RESULT 15
US-09-620-412C-25/c
: Sequence 25, Application US/09620412C
: Patent No. 6448234
: GENERAL INFORMATION:
: APPLICANT: Steven P. Fling
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
: TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
: FILE REFERENCE: 210121.469C7
: CURRENT APPLICATION NUMBER: US/09/620,412C
: CURRENT FILING DATE: 2000-07-20
: NUMBER OF SEQ ID NOS: 363
: SOFTWARE: FastSeq for Windows Version 3.0/4.0
: SEQ ID NO 25
: LENGTH: 696
: TYPE: DNA
: ORGANISM: Chlamydia trachomatis
US-09-620-412C-25

```

```

Query Match          9.1%; Score 16; DB 4; Length 696;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 117 AAGGACACGACAGTA 132
    |||||||
Db 538 AAGGACACGACAGTA 523

```

Search completed: July 28, 2003, 15:58:55  
 Job time : 10.5631 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 14:30:21 ; Search time 43.097 Seconds  
(Without alignments)  
8377.033 Million cell updates/sec

Title: US-10-010-408-1\_COPY\_1534\_1708  
Perfect score: 175  
Sequence: 1 AGTCACGAGACTTGAGCTT.....GCCTAGATTAACACCCCAA 175

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 1439767 seqs, 1031500376 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2879534

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Published Applications\_NA:\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/2/pubpna/PCr\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/2/pubpna/PCrUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
10: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq1:\*  
11: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq2:\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq3:\*  
13: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
15: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*  
16: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*  
17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query Match | Length  | ID                      | Description          |
|------------|-------|-------------|---------|-------------------------|----------------------|
| 1          | 175   | 100.0       | 1708    | 15 US-10-010-408-1      | Sequence 1, Appl1    |
| 2          | 121   | 69.1        | 439     | 12 US-09-956-622A-23    | Sequence 23, Appl1   |
| 3          | 35    | 20.0        | 1734    | 15 US-10-112-267-17     | Sequence 17, Appl1   |
| 4          | 35    | 20.0        | 1734    | 15 US-10-112-267-18     | Sequence 18, Appl1   |
| 5          | 18    | 10.3        | 1671    | 10 US-09-841-132-378    | Sequence 378, Appl1  |
| 6          | 18    | 10.3        | 2586    | 10 US-09-815-242-9990   | Sequence 9990, App   |
| 7          | 18    | 10.3        | 3073    | 15 US-10-027-632-114289 | Sequence 114289, App |
| 8          | 18    | 10.3        | 3073    | 15 US-10-027-632-114290 | Sequence 114290, App |
| 9          | 18    | 10.3        | 3186778 | 15 US-10-027-632-174961 | Sequence 174961, App |
| 10         | 17    | 9.7         | 24      | 15 US-10-112-267-110    | Sequence 110, App    |
| 11         | 17    | 9.7         | 141     | 11 US-09-783-590-7073   | Sequence 7073, App   |
| 12         | 17    | 9.7         | 451     | 12 US-09-918-995-12392  | Sequence 12392, App  |
| 13         | 17    | 9.7         | 627     | 15 US-10-027-632-204519 | Sequence 204519, App |
| 14         | 17    | 9.7         | 627     | 15 US-10-027-632-204520 | Sequence 204520, App |
| 15         | 17    | 9.7         | 627     | 15 US-10-027-632-204521 | Sequence 204521, App |
| 16         | 17    | 9.7         | 738     | 15 US-10-027-632-32541  | Sequence 32541, App  |

|      |    |     |      |                         |                      |
|------|----|-----|------|-------------------------|----------------------|
| C 17 | 17 | 9.7 | 742  | 12 US-09-968-433-2      | Sequence 2, Appl1    |
| 18   | 17 | 9.7 | 774  | 15 US-10-027-632-114990 | Sequence 114990, App |
| 19   | 17 | 9.7 | 774  | 15 US-10-027-632-114991 | Sequence 114991, App |
| 20   | 17 | 9.7 | 940  | 15 US-10-027-632-261119 | Sequence 261119, App |
| 21   | 17 | 9.7 | 940  | 15 US-10-027-632-261120 | Sequence 261120, App |
| 22   | 17 | 9.7 | 2126 | 15 US-10-052-586-495    | Sequence 495, App    |
| 23   | 17 | 9.7 | 2126 | 15 US-10-174-758-495    | Sequence 495, App    |
| 24   | 17 | 9.7 | 2126 | 15 US-10-176-758-495    | Sequence 495, App    |
| 25   | 17 | 9.7 | 2126 | 15 US-10-175-737-495    | Sequence 495, App    |
| 26   | 17 | 9.7 | 2126 | 15 US-10-173-706-495    | Sequence 495, App    |
| 27   | 17 | 9.7 | 2126 | 15 US-10-175-738-495    | Sequence 495, App    |
| 28   | 17 | 9.7 | 2126 | 15 US-10-175-752-495    | Sequence 495, App    |
| 29   | 17 | 9.7 | 2126 | 15 US-10-176-482-495    | Sequence 495, App    |
| 30   | 17 | 9.7 | 2126 | 15 US-10-176-757-495    | Sequence 495, App    |
| 31   | 17 | 9.7 | 2126 | 15 US-10-176-913-495    | Sequence 495, App    |
| 32   | 17 | 9.7 | 2126 | 15 US-10-180-552-495    | Sequence 495, App    |
| 33   | 17 | 9.7 | 2126 | 15 US-10-180-557-495    | Sequence 495, App    |
| 34   | 17 | 9.7 | 2126 | 15 US-10-173-700-495    | Sequence 495, App    |
| 35   | 17 | 9.7 | 2126 | 15 US-10-174-572-495    | Sequence 495, App    |
| 36   | 17 | 9.7 | 2126 | 15 US-10-174-579-495    | Sequence 495, App    |
| 37   | 17 | 9.7 | 2126 | 15 US-10-174-582-495    | Sequence 495, App    |
| 38   | 17 | 9.7 | 2126 | 15 US-10-174-588-495    | Sequence 495, App    |
| 39   | 17 | 9.7 | 2126 | 15 US-10-175-739-495    | Sequence 495, App    |
| 40   | 17 | 9.7 | 2126 | 15 US-10-175-740-495    | Sequence 495, App    |
| 41   | 17 | 9.7 | 2126 | 15 US-10-175-743-495    | Sequence 495, App    |
| 42   | 17 | 9.7 | 2126 | 15 US-10-176-488-495    | Sequence 495, App    |
| 43   | 17 | 9.7 | 2126 | 15 US-10-176-492-495    | Sequence 495, App    |
| 44   | 17 | 9.7 | 2126 | 15 US-10-176-747-495    | Sequence 495, App    |
| 45   | 17 | 9.7 | 2126 | 15 US-10-176-750-495    | Sequence 495, App    |

# ALIGNMENTS

RESULT 1  
US-10-010-408-1  
Sequence 1, Application US/10010408  
Publication No. US20020165185A1

GENERAL INFORMATION:  
APPLICANT: John J. Castellot, Jr.  
TITLE OF INVENTION: NO. US20020165185A1 Heparin-Induced CCN-Like Molecules and Uses Therefor

NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD, LLP  
STREET: 28 State Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109

COMPUTER READABLE FORM:  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/010,408  
FILING DATE: 07-Dec-2001  
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/044,273  
FILING DATE: March 19, 1998  
APPLICATION NUMBER: <Unknown>  
FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:  
NAME: Amy E. Mandragouras  
REGISTRATION NUMBER: 36,207  
REFERENCE/DOCKET NUMBER: MBI-004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)742-4214  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:

LENGTH: 1708 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 249..1001  
SEQUENCE DESCRIPTION: SEQ ID NO: 1  
US-10-010-408-1

Query Match 100.0%; Score 175; DB 15; Length 1708;  
Best Local Similarity 100.0%; Pred. No. 3.1e-85  
Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGTCAGGAACCTTGACTTGTATTTTCAGGAATGCACATCTTAAAGCACTGCAAAAC 60  
1534 AGTCAGGAACCTTGACTTGTATTTTCAGGAATGCACATCTTAAAGCACTGCAAAAC 1593  
DB 61 AGGAGGCTCCACACTCTGGCAGGCCAGGCCCTTCTGTCAGCATGAGAAAGACAAG 120  
1594 AGGAGGCTCCACACTCTGGCAGGCCAGGCCCTTCTGTCAGCATGAGAAAGACAAG 1653  
QY 121 GACAGAGAGTACTCTCTCTGGAGAGACTAGTACCTTGAATTAACCCCAA 175  
DB 1654 GACAGAGAGTACTCTCTCTGGAGAGACTAGTACCTTGAATTAACCCCAA 1708

## RESULT 2

US-09-956-622A-23  
Sequence 23, Application US/09956622A  
Publication No. US20030091973A1  
GENERAL INFORMATION:  
APPLICANT: Horesovsky, Gregory J  
APPLICANT: No. US20030091973A1 II, L. Staton  
APPLICANT: Raha, Debashish  
TITLE OF INVENTION: Method of Identifying Osteoregenerative Agents Using  
FILE REFERENCE: 21402-445  
CURRENT APPLICATION NUMBER: US/09/956,622A  
CURRENT FILING DATE: 2001-09-19  
PRIOR APPLICATION NUMBER: 60/233,579  
PRIOR FILING DATE: 2000-09-19  
NUMBER OF SEQ ID NOS: 53  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 23  
LENGTH: 439  
TYPE: DNA  
ORGANISM: Rattus norvegicus  
US-09-956-622A-23

Query Match 69.1%; Score 121; DB 12; Length 439;  
Best Local Similarity 100.0%; Pred. No. 7.4e-56  
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGTCAGGAACCTTGACTTGTATTTTCAGGAATGCACATCTTAAAGCACTGCAAAAC 60  
DB 285 AGTCAGGAACCTTGACTTGTATTTTCAGGAATGCACATCTTAAAGCACTGCAAAAC 344  
QY 61 AGGAGGCTCCACACTCTGGCAGGCCAGGCCCTTCTGTCAGCATGAGAAAGACAAG 120  
DB 345 AGGAGGCTCCACACTCTGGCAGGCCAGGCCCTTCTGTCAGCATGAGAAAGACAAG 404  
QY 121 G 121  
DB 405 G 405

RESULT 3  
US-10-112-267-17  
Sequence 17, Application US/10112267  
Publication No. US20030068678A1  
GENERAL INFORMATION:

APPLICANT: Botstein, David A.  
APPLICANT: Cohen, Robert  
APPLICANT: Goddard, Audrey  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: Lawrence, David A.  
APPLICANT: Levine, Arnold J.  
APPLICANT: Pennica, Diane  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME  
FILE REFERENCE: P11762  
CURRENT APPLICATION NUMBER: US/10/112,267  
CURRENT FILING DATE: 2002-03-27  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/182,145B  
PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-29  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/063,704  
PRIOR FILING DATE: EARLIER FILING DATE: 1997-10-29  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/073,612  
PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-04  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/081,695  
PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-14  
NUMBER OF SEQ ID NOS: 156  
SEQ ID NO 17  
LENGTH: 1734  
TYPE: DNA  
ORGANISM: Mus musculus  
US-10-112-267-17

Query Match 20.0%; Score 35; DB 15; Length 1734;  
Best Local Similarity 100.0%; Pred. No. 4e-09;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 TCTGGCAGGCGCAGGCGCTTCTCTCAGCATGAGA 111  
DB 1611 TCTGGCAGGCGCAGGCGCTTCTCTCAGCATGAGA 1645

## RESULT 4

US-10-112-267-18/C  
Sequence 18, Application US/10112267  
Publication No. US20030068678A1  
GENERAL INFORMATION:  
APPLICANT: Botstein, David A.  
APPLICANT: Cohen, Robert  
APPLICANT: Goddard, Audrey  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: Lawrence, David A.  
APPLICANT: Levine, Arnold J.  
APPLICANT: Pennica, Diane  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME  
FILE REFERENCE: P11762  
CURRENT APPLICATION NUMBER: US/10/112,267  
CURRENT FILING DATE: 2002-03-27  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/182,145B  
PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-29  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/063,704  
PRIOR FILING DATE: EARLIER FILING DATE: 1997-10-29  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/073,612  
PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-04  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/081,695  
PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-14  
NUMBER OF SEQ ID NOS: 156  
SEQ ID NO 18  
LENGTH: 1734  
TYPE: DNA  
ORGANISM: Mus musculus  
US-10-112-267-18

Query Match 20.0%; Score 35; DB 15; Length 1734;

Best Local Similarity 100.0%; Pred. No. 4e-09;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 77 TCTGGCAGCCGCGCTTCTCTTCACGATGACA 111  
Db 124 TCTGGCAGCCGCGCTTCTCTTCACGATGACA 90

RESULT 5  
US-09-841-132-378/C

Sequence 378, Application US/09841132  
Patent No. US20020061848A1  
GENERAL INFORMATION:  
APPLICANT: Bhatia, Ajay  
APPLICANT: Shetla, Yashir A.W.  
APPLICANT: Probst, Peter  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND  
FILE REFERENCE: 210121.469C8  
CURRENT APPLICATION NUMBER: US/09/841.132  
CURRENT FILING DATE: 2001-04-23  
NUMBER OF SEQ ID NOS: 599  
SOFTWARE: FastSeq for Windows Version 3.0/4.0  
SEQ ID NO 378  
LENGTH: 1671  
TYPE: DNA  
ORGANISM: Chlamydia pneumoniae  
US-09-841-132-378.

Query Match 10.3%; Score 18; DB 10; Length 1671;  
Best Local Similarity 100.0%; Pred. No. 7;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 TCCGGAAGCTGAGCTTT 20  
Db 1227 TCCGGAAGCTGAGCTTT 1210

RESULT 6  
US-09-815-242-9990

Sequence 9990, Application US/09815242  
Patent No. US20020061369A1  
GENERAL INFORMATION:  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlson, Karl L.  
APPLICANT: Zyskind, Judith W.  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John D.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
APPLICANT: Xu, H. Howard  
TITLE OF INVENTION: Identification of Essential Genes in  
FILE REFERENCE: ELITRA.011A  
CURRENT APPLICATION NUMBER: US/09/815.242  
CURRENT FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 14110  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 9990

LENGTH: 2586  
TYPE: DNA  
ORGANISM: Salmonella typhi  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)...(2586)  
NAME/KEY: misc-feature  
LOCATION: (1)...(2586)  
OTHER INFORMATION: n = A,T,C or G  
US-09-815-242-9990

Query Match 10.3%; Score 18; DB 10; Length 2586;  
Best Local Similarity 100.0%; Pred. No. 6.8;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 51 CTCGCAAAACAGCAGGC 68  
Db 808 CTCGCAAAACAGCAGGC 825

RESULT 7  
US-10-027-632-114289

Sequence 114289, Application US/10027632  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Polymorphisms in the Human Genome  
FILE REFERENCE: 108827.129  
CURRENT APPLICATION NUMBER: US/10/027.632  
CURRENT FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/218,006  
PRIOR FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 60/198,676  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: US 60/193,483  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: US 60/185,218  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/167,363  
PRIOR FILING DATE: 1999-11-23  
PRIOR APPLICATION NUMBER: US 60/156,358  
PRIOR FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: US 60/146,002  
PRIOR FILING DATE: 1999-08-09  
NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 114289  
LENGTH: 3073  
TYPE: DNA  
ORGANISM: Human  
US-10-027-632-114289

Query Match 10.3%; Score 18; DB 15; Length 3073;  
Best Local Similarity 100.0%; Pred. No. 6.8;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 64 AAGCTCCACACCTCTGG 81  
Db 2440 AAGCTCCACACCTCTGG 2457

RESULT 8  
US-10-027-632-114290

Sequence 114290, Application US/10027632  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
FILE REFERENCE: 108827.129  
CURRENT APPLICATION NUMBER: US/10/027.632  
CURRENT FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/218,006  
PRIOR FILING DATE: 2000-07-12

```

: PRIOR APPLICATION NUMBER: US 60/198,676
: PRIOR FILING DATE: 2000-04-20
: PRIOR APPLICATION NUMBER: US 60/193,483
: PRIOR FILING DATE: 2000-03-29
: PRIOR APPLICATION NUMBER: US 60/185,218
: PRIOR FILING DATE: 2000-02-24
: PRIOR APPLICATION NUMBER: US 60/167,363
: PRIOR FILING DATE: 1999-11-23
: PRIOR APPLICATION NUMBER: US 60/156,358
: PRIOR FILING DATE: 1999-09-28
: PRIOR APPLICATION NUMBER: US 60/146,002
: PRIOR FILING DATE: 1999-08-09
: NUMBER OF SEQ ID NOS: 325720
: SOFTWARE: fastseq for windows version 4.0
: SEQ ID NO 114290
: LENGTH: 3073
: TYPE: DNA
: ORGANISM: Human
: US-10-027-632-114290
```

```
Query Match
Best Local Similarity 100.0%; Score 18; DB 15; Length 3073;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 64 AAGGCTCCACACTCTG 81
Db 2440 AAGGCTCCACACTCTG 2457
```

```
RESULT 9
US-10-027-632-174961
: Sequence 174961, Application US/10027632
```

```

: GENERAL INFORMATION:
: APPLICANT: Wang, David G.
: TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
: FILE REFERENCE: 108827.129
: CURRENT APPLICATION NUMBER: US/10/027,632
: CURRENT FILING DATE: 2002-04-30
: PRIOR APPLICATION NUMBER: US 60/218,006
: PRIOR FILING DATE: 2000-07-12
: PRIOR APPLICATION NUMBER: US 60/198,676
: PRIOR FILING DATE: 2000-04-20
: PRIOR APPLICATION NUMBER: US 60/193,483
: PRIOR FILING DATE: 2000-03-29
: PRIOR APPLICATION NUMBER: US 60/185,218
: PRIOR FILING DATE: 2000-02-24
: PRIOR APPLICATION NUMBER: US 60/167,363
: PRIOR FILING DATE: 1999-11-23
: PRIOR APPLICATION NUMBER: US 60/156,358
: PRIOR FILING DATE: 1999-09-28
: PRIOR APPLICATION NUMBER: US 60/146,002
: PRIOR FILING DATE: 1999-08-09
: NUMBER OF SEQ ID NOS: 325720
: SOFTWARE: fastseq for windows version 4.0
: SEQ ID NO 174961
: LENGTH: 3186778
: TYPE: DNA
: ORGANISM: Human
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1)...(3186778)
: OTHER INFORMATION: n = A,T,C or G
: US-10-027-632-174961
```

```
Query Match
Best Local Similarity 100.0%; Score 18; DB 15; Length 3186778;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 47 AGCACTCCGCAAAACAGA 64
Db 805006 AGCACTCCGCAAAACAGA 805023
```

```
RESULT 10
US-10-112-267-110/c
: Sequence 110, Application US/10112267
: Publication No. US20030068678A1
```

```

: GENERAL INFORMATION:
: APPLICANT: Botstein, David A.
: APPLICANT: Cohen, Robert
: APPLICANT: Goddard, Audrey
: APPLICANT: Gurney, Austin L.
: APPLICANT: Hillan, Kenneth J.
: APPLICANT: Lawrence, David A.
: APPLICANT: Levine, Arnold J.
: APPLICANT: Pennica, Diane
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Wood, William I.
: TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
: FILE REFERENCE: P11762
: CURRENT APPLICATION NUMBER: US/10/112,267
: CURRENT FILING DATE: 2002-03-27
: PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/183,145B
: PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-29
: PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/063,704
: PRIOR FILING DATE: EARLIER FILING DATE: 1997-10-29
: PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/073,612
: PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-04
: PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/081,695
: PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-14
: NUMBER OF SEQ ID NOS: 156
: SEQ ID NO 110
: LENGTH: 24
: TYPE: DNA
: ORGANISM: Artificial sequence
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 1-24
: OTHER INFORMATION: Sequence is synthesized.
: US-10-112-267-110
```

```
Query Match
Best Local Similarity 9.7%; Score 17; DB 15; Length 24;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 77 TCTGGCAGCGCAGGCC 93
Db 17 TCTGGCAGCGCAGGCC 1
```

```
RESULT 11
US-09-783-590-7073
: Sequence 7073, Application US/09783590
: Patent No. US20020110850A1
```

```

: GENERAL INFORMATION:
: APPLICANT: Dillion, Patrick J.
: APPLICANT: Haseltine, William A.
: APPLICANT: Li, Haodong
: APPLICANT: Rosen, Craig A.
: APPLICANT: Ruben, Steven M.
: TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
: FILE REFERENCE: PO-16 2C1
: CURRENT APPLICATION NUMBER: US/09/783,590
: CURRENT FILING DATE: 2000-02-15
: PRIOR APPLICATION NUMBER: 08/420,856
: PRIOR FILING DATE: 1995-04-12
: PRIOR APPLICATION NUMBER: 08/346,731
: PRIOR FILING DATE: 1994-11-21
: NUMBER OF SEQ ID NOS: 12485
: SOFTWARE: Patentin Ver. 2.0
: SEQ ID NO 7073
: LENGTH: 141
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
```



; NAME/KEY: misc feature  
; LOCATION: (2)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc feature  
; LOCATION: (67)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc feature  
; LOCATION: (70)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc feature  
; LOCATION: (112)  
; OTHER INFORMATION: n equals a,t,g, or c  
US-09-783-590-7073

Query Match 9.7%; Score 17; DB 11; Length 141;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 95 TTCTCTTCAGCATGAGA 111  
|||||  
Db 90 TTCTCTTCAGCATGAGA 106

RESULT 12  
US-09-918-995-12392/C  
; Sequence 12392, Application US/09918995  
; Publication No. US20030073623A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc.  
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED  
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES  
; FILE REFERENCE: 20411-756  
; CURRENT APPLICATION NUMBER: US/09/918,995  
; PRIOR FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: US/09/235,076  
; PRIOR FILING DATE: 1999-01-20  
; NUMBER OF SEQ ID NOS: 38054  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 12392  
; LENGTH: 451  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)...(451)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-918-995-12392

Query Match 9.7%; Score 17; DB 12; Length 451;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TCCAGGAAGTGGAGCTT 19  
|||||  
Db 174 TCCAGGAAGTGGAGCTT 158

RESULT 13  
US-10-027-632-204519/C  
; Sequence 204519, Application US/10027632  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; TITLE OF INVENTION: Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; PRIOR FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29

; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 204519  
; LENGTH: 627  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-204519

Query Match 9.7%; Score 17; DB 15; Length 627;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 ATGAGAAAGACAGGGA 122  
|||||  
Db 457 ATGAGAAAGACAGGGA 441

RESULT 14  
US-10-027-632-204520/C  
; Sequence 204520, Application US/10027632  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; TITLE OF INVENTION: Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; PRIOR FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 204520  
; LENGTH: 627  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-204520

Query Match 9.7%; Score 17; DB 15; Length 627;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 ATGAGAAAGACAGGGA 122  
|||||  
Db 457 ATGAGAAAGACAGGGA 441

RESULT 15  
US-10-027-632-204521/C  
; Sequence 204521, Application US/10027632  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; TITLE OF INVENTION: Polymorphisms in the Human Genome

FILE REFERENCE: 108827.129  
CURRENT APPLICATION NUMBER: US/10/027,632  
CURRENT FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/218,006  
PRIOR FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 60/198,676  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: US 60/193,483  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: US 60/185,218  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/167,363  
PRIOR FILING DATE: 1999-11-23  
PRIOR APPLICATION NUMBER: US 60/156,358  
PRIOR FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: US 60/146,002  
PRIOR FILING DATE: 1999-08-09  
NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 204521  
LENGTH: 627  
TYPE: DNA  
ORGANISM: Human  
US-10-027-632-204521

Query Match 9.7%; Score 17; DB 15; Length 627;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 ATGAGAAAGACAAGGA 122  
DB 457 ATGAGAAAGACAAGGA 441

Search completed: July 28, 2003, 15:36:37  
Job time : 48.097 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 15:16:26 ; Search time 310.442 Seconds  
(without alignments)  
9129.604 Million cell updates/sec

Title: US-10-010-408-1\_COPY\_1534\_1708

Perfect score: 175

Sequence: 1 AGTCCAGGAGACTTGACTT.....GCCATGATTAACACCAAA 175

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size : 0

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estrov:\*  
6: em\_estrpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description        |
|------------|-------|-------------|--------|----|--------------------|
| 1          | 153   | 87.4        | 762    | 14 | B0195526 UI-R-CNI- |
| 2          | 35    | 20.0        | 315    | 10 | B8374499 BB374499  |
| 3          | 35    | 20.0        | 369    | 10 | BB792544 BB792544  |
| 4          | 35    | 20.0        | 418    | 9  | AT225477 u88B01.Y  |
| 5          | 35    | 20.0        | 792    | 12 | BF138093 B01785003 |
| 6          | 22    | 12.6        | 307    | 10 | BB220676 BB220676  |

|   |    |    |      |      |    |          |                    |
|---|----|----|------|------|----|----------|--------------------|
| C | 7  | 20 | 11.4 | 299  | 12 | BF903057 | BF903057 CM3-MT019 |
|   | 8  | 20 | 11.4 | 313  | 10 | BB367824 | BB367824 BB367824  |
|   | 9  | 20 | 11.4 | 538  | 17 | A2372255 | A2372255 IM0124T03 |
|   | 10 | 20 | 11.4 | 683  | 17 | AG078128 | AG078128 Pan trogl |
|   | 11 | 19 | 10.9 | 546  | 10 | AM606058 | AM606058 99634 MAR |
|   | 12 | 19 | 10.9 | 591  | 17 | BH112293 | BH112293 RPT-24-2  |
|   | 13 | 19 | 10.9 | 932  | 12 | BE778127 | BE778127 601463356 |
|   | 14 | 19 | 10.9 | 957  | 13 | BI334722 | BI334722 60299615  |
|   | 15 | 19 | 10.9 | 1146 | 17 | CMS06NTV | AL406865 T3 end of |
|   | 16 | 19 | 10.9 | 5332 | 17 | AF101910 | AF101910 AF101910  |
|   | 17 | 18 | 10.3 | 108  | 12 | BE809197 | BE809197 214594 MA |
|   | 18 | 18 | 10.3 | 111  | 10 | AM352985 | AM352985 35259 MAR |
|   | 19 | 18 | 10.3 | 281  | 10 | BB309266 | BB309266 BB309266  |
|   | 20 | 18 | 10.3 | 297  | 14 | F14567   | F14567 SSC2H04 Por |
|   | 21 | 18 | 10.3 | 320  | 12 | BF561635 | BF561635 UI-R-CO-1 |
|   | 22 | 18 | 10.3 | 367  | 9  | AA804915 | AA804915 OA89D02.S |
|   | 23 | 18 | 10.3 | 371  | 10 | AV667415 | AV667415 AV667415  |
|   | 24 | 18 | 10.3 | 375  | 17 | A0134190 | A0134190 HS_3057_A |
|   | 25 | 18 | 10.3 | 383  | 13 | BI327290 | BI327290 AR076E021 |
|   | 26 | 18 | 10.3 | 385  | 12 | BF193668 | BF193668 245334 MA |
|   | 27 | 18 | 10.3 | 386  | 18 | AM654343 | AM654343 103751 MA |
|   | 28 | 18 | 10.3 | 389  | 17 | A2277972 | A2277972 RPT-23-1  |
|   | 29 | 18 | 10.3 | 395  | 10 | AM977810 | AM977810 EST389814 |
|   | 30 | 18 | 10.3 | 426  | 9  | AA767711 | AA767711 OB49B08.S |
|   | 31 | 18 | 10.3 | 457  | 10 | AM776806 | AM776806 EST335871 |
|   | 32 | 18 | 10.3 | 461  | 17 | AZ838436 | AZ838436 2M0134E10 |
|   | 33 | 18 | 10.3 | 482  | 12 | BF010204 | BF010204 S9508.Y   |
|   | 34 | 18 | 10.3 | 487  | 12 | BF806741 | BF806741 IL5-C1014 |
|   | 35 | 18 | 10.3 | 491  | 14 | BM703250 | BM703250 UI-E-C11- |
|   | 36 | 18 | 10.3 | 491  | 9  | AU083781 | AU083781 AU083781  |
|   | 37 | 18 | 10.3 | 491  | 17 | A2965530 | A2965530 2M0235J07 |
|   | 38 | 18 | 10.3 | 508  | 17 | AQ430430 | AQ430430 HS_5095_A |
|   | 39 | 18 | 10.3 | 519  | 10 | BE203038 | BE203038 EST403060 |
|   | 40 | 18 | 10.3 | 530  | 13 | BJ538294 | BJ538294 BJS38294  |
|   | 41 | 18 | 10.3 | 534  | 13 | BJ467272 | BJ467272 BJ467272  |
|   | 42 | 18 | 10.3 | 543  | 10 | AM142440 | AM142440 EST292692 |
|   | 43 | 18 | 10.3 | 560  | 14 | BQ201661 | BQ201661 UI-R-D01- |
|   | 44 | 18 | 10.3 | 564  | 10 | BE342161 | BE342161 EST395003 |
|   | 45 | 18 | 10.3 | 627  | 17 | AZ838449 | AZ838449 2M0134G11 |

#### ALIGNMENTS

RESULT 1  
B0195526/c 762 bp mRNA linear EST 30-APR-2002  
LOCUS UI-R-CNI-cmq-k-07-0-UI.s1 UI-R-CNI Rattus norvegicus cDNA clone  
DEFINITION UI-R-CNI-cmq-k-07-0-UI 3', mRNA sequence.  
ACCESSION B0195526  
VERSION B0195526.1 GI:20371077  
KEYWORDS EST.  
SOURCE Norway rat.  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae;  
Rattus.  
REFERENCE 1 (bases 1 to 762)  
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.  
TITLE Normalization and subtraction: two approaches to facilitate gene  
discovery  
JOURNAL Genome Res. 6 (9), 791-806 (1996)  
MEDLINE 97044477  
COMMENT Contact: Soares, MB  
Program for Rat Gene Discovery and Mapping  
University of Iowa  
451 Eckstein Medical Research Building Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: msquares@blue.weeg.uiowa.edu  
The sequence contained an oligo-dT track that was present in the  
oligonucleotide that was used to prime the synthesis of first  
strand cDNA and therefore this may represent a bonafide poly A

tall. The sequence tag present in the cDNA between the NotI site and the oligo-OT track served to identify it as a clone from the normalized cervix library cDNA Library Preparation: M.B. Soares Lab Genetics (www.resgen.com)

Seq primer: M13 Forward

PolyA-Tes

# FEATURES

## source

Location/Qualifiers

1. 762

/organism="Rattus norvegicus"

/strain="Sprague-Dawley"

/db\_xref="taxon:10116"

/clone="UI-R-CNI-cmq-x-07-0-UI"

/clone\_lib="UI-R-CNI"

/dev\_stage="adult"

/lab\_host="DH10B (Life Technologies)"

/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker. Site 1: Not I; Site 2: Eco RI; The UI-R-CNI library is a subtracted library derived from the following pool of seven normalized rat libraries: normalized rat seminal vesicles, normalized rat penis, normalized rat bladder, normalized rat cervix, normalized rat brown adipose, normalized rat fundus, and normalized rat salivary gland. It was constructed according to the procedure described by Bonaldo, Lennon & Soares (Genome Research Genome 6: 791-806, 1996). For construction of the CNI library, plasmid DNA from the pool of normalized libraries was electroporated into competent bacteria for the production of single-stranded circular DNA. This was then used as a tracer in a subtractive hybridization with a driver (PCR amplified inserts from a plasmid DNA template preparation) comprising: a) a pool of about 34,000 clones from the Rat Urogen Set corresponding to plates R-5-AA-NN excluding plates R-5-MM and MN. This pool represented 40% of the final driver population. b) a pool of about 29,000 clones from subtracted libraries: CNO and CNI corresponding to plates R-CAO-AMV through R-CAO-AXS, R-CAO-AZX through R-CAO-BAZ, R-CAO-BBE through R-CAO-BHY, R-CAO-BJS, R-CAO-BKE, R-CAO-BKG-H, R-CAO-BKJ-K, R-CAO-BKP through R-CAO-BKS, R-CAO-BKV-V, R-CAO-BLY through R-CAO-BMA, R-CAO-BMC through R-CAO-BME, R-CAO-BNS, R-CAO-BOB through R-CAO-BOJ, R-CAO-BPA through R-CAO-BPG, R-CAI-BBA through R-CAI-BDA, R-CAI-BHZ through R-CAI-BTF, R-CAI-BUR, R-CAI-BUT through R-CAI-BKB, R-CAI-BKD, R-CAI-BKF, R-CAI-BKI, R-CAI-BKT, R-CAI-BLF, R-CAI-BLR through R-CAI-BLN, R-CAI-BLS, R-CAI-BLV, R-CAI-BMR, and R-CAI-BLE. The resulting pool represented 20% of the final driver population. c) a pool of about 15,000 clones from non-normalized libraries CS0s, CT0s, CU0s, CM0s, CX0s and normalized libraries CS0, CT0, CU0, CM0, and CX0 corresponding to plates R-CS0s-CBD through R-CS0s-CBO, R-CT0s-CAM through R-CT0s-CAX, R-CU0s-CBP through R-CU0s-CCA, R-CM0s-CCB through R-CM0s-CCM, R-CX0s-CCN through R-CX0s-CCX, R-CS0-BSD, R-CS0-BMD through R-CS0-BTV, R-CS0-BVM, R-CT0-BTW through R-CT0-BDP, R-CT0-BVN, R-CU0-BUQ through R-CU0-BVL, R-CM0-BVJ through R-CM0-BWP, R-CX0-BXN through R-CX0-BXO, R-CX0-BMO through R-CX0-BXM. The resulting pool represented 5% of the final driver population. d) a pool of about 5,000 clones (1,000 from non-normalized eye library CV0 and 4,000 from normalized eye library CV1) corresponding to plates R-CV0-BBR through R-CV0-BRR, R-CV1-BBS through R-CV1-BSC, R-CV1-BSE through R-CV1-BTC, and R-CV1-BVO through R-CV1-BVU. This pool represented about 5% of the final driver population. e) A pool of about 10,000 clones from subtracted library BS2, BV0 and BV0P (7-9.5 kb cDNA library fraction from rat whole embryo), and BX0 (0.5-7kb cDNA library fraction from rat whole embryo) corresponding to plates R-BS2-BDB through R-BS2-BPB, R-BV0-ANR through R-BV0-AOI through R-BV0-AOX, and R-BX0-AGY through R-BX0-ASH. The resulting pool represented 5% of the final driver population. f) a pool of about 7,000 clones from the seven non-normalized libraries that make up the tracer

BASE COUNT 177 a 176 c 198 g 211 t

## ORIGIN

Query Match 87.4%; Score 153; DB 14; Length 762;  
Best Local Similarity 100.0%; Pred. No. 5.7e-70;

Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 AGTCGAGGAACTGAGCTTGTATTTTCAGAGTGCATCTCTTAAGCACTGCAAAAC 60

202 AGTCGAGGAACTGAGCTTGTATTTTCAGAGTGCATCTCTTAAGCACTGCAAAAC 143

61 AGGAGGCTCCACACCTGCGAGGCGCCCTTCTCTTCAGATGAGAAAGACAGG 120

142 AGGAGGCTCCACACCTGCGAGGCGCCCTTCTCTTCAGATGAGAAAGACAGG 83

121 GACAGCAGAGTACTCTCTCTGAGACATGATC 153

82 GACAGCAGAGTACTCTCTCTGAGACATGATC 50

RESULT 2  
LOCUS BB374499 315 bp mRNA linear EST 13-JUL-2000

DEFINITION BB374499 RIKEN full-length enriched, 16 days embryo head Mus musculus cDNA clone C130075616 3' similar to AFI26063 Mus musculus connective tissue growth factor-like protein precursor (Ctgfl) mRNA

ACCESSION BB374499  
VERSION BB374499.1 GI:9086993

KEYWORDS house mouse.  
SOURCE Mus musculus.  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 315)  
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

Komno, H., Aizawa, K., Akahira, S., Akiyama, J., Arikawa, T., Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hirozane, T., Horii, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M., Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N., Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurihara, C., Kusakabe, M.

```

cdna_lib="Riken full-length enriched, kidney CCL-142 RAG"
/tissue_type="kidney"
/cell_line="CCL-142 RAG"
/note="pooled cell lines : (cell_line=CRJ-1751 WEHI 164),
(cell_line=CRJ-2116 JC), (cell_line=RCB-0035 WEHI-3),
(cell_line=RCB-0464 Meth-1), (cell_line=RCB-0545 OHTA),
(cell_line=RCB-0559 K-1, FI), (cell_line=RCB-1283 B16
melanoma), (cell_line=RCB-1702 WEHI 231
), (cell_line=Leydig cells, cell_line=CRJ-2065 MLTc-1),
(cell_type=Nullipotent stem cell, cell_line=CRJ-2070 NE),
(tissue_type=bladder, cell_line=RCB-0544 MBF-2),
(tissue_type=bone marrow, cell_type=stroma cell,
cell_line=CRJ-2028 SR-4987), (tissue_type=colon,
cell_line=RCB-0549 C1e-H3), (tissue_type=kidney,

```

cell\_line=CCl-142 RAG), (tissue\_type=submandibular gland, cell\_line=CRL-1734 SCA-9 clone 15), (strain=BALB/C, cell\_type=B cells, cell\_line=CRL-1669 BCL1 Clone 13-20-3B3), (strain=C3H, tissue\_type=brain, cell\_line=CRL-1443 BC3H1)"

BASE COUNT 102 a 90 c 86 g 91 t

Query Match 20.0%; Score 35; DB 10; Length 369;  
Best Local Similarity 100.0%; Pred. No. 1.1e-07;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 TCTGGCAGCCAGCGCCCTTCTCTCAGCATGAGA 111  
|||||  
Db 262 TCTGGCAGCCAGCGCCCTTCTCTCAGCATGAGA 296

RESULT 4  
AI225477 418 bp mRNA linear EST 29-OCT-1998  
LOCUS ue88B01.y1 Soares\_NMPu Mus musculus cDNA clone IMAGE:1498153 5',  
DEFINITION mRNA sequence.  
ACCESSION AI225477 GI:3808530  
VERSION AI225477.1  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus.

REFERENCE  
AUTHORS Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 418)  
Marras, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, B., Schellenberg, K., Steptoe, M., Tan, F., Underwood, R., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.

TITLE The WashU-HMI Mouse EST Project  
JOURNAL Unpublished (1996)  
COMMENT Contact: Maira M/Mouse EST Project  
WashU-HMI Mouse EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mouseest@watson.wustl.edu  
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.  
MGI:935757  
Seq primer: -40RP from Glibco  
High quality sequence stop: 403.  
Location/Qualifiers

FEATURES  
source  
1. 418  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone\_image="IMAGE:1498153"  
/clone\_lib="Soares\_NMPu"  
/sex="female"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/note="Organ: uterus; Vector: pMT73D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from pregnant mouse uterus, and was then primed with a Not I - Oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pMT73 vector. Library is normalized. Library was constructed by Bento Soares and M. Fatima Bonafido."

BASE COUNT 121 a 98 c 100 g 99 t

Query Match 20.0%; Score 35; DB 9; Length 418;  
Best Local Similarity 100.0%; Pred. No. 1.2e-07;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 TCTGGCAGCCAGCGCCCTTCTCTCAGCATGAGA 111  
|||||  
Db 300 TCTGGCAGCCAGCGCCCTTCTCTCAGCATGAGA 334

RESULT 5  
BF138093 792 bp mRNA linear EST 24-OCT-2000  
LOCUS 601785003F1 NCI\_CGAP\_Lu30 Mus musculus cDNA clone IMAGE:4012801 5',  
DEFINITION mRNA sequence.  
ACCESSION BF138093  
VERSION BF138093.1 GI:10977133  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus.

REFERENCE  
AUTHORS Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 792)  
NIH-MGC http://mgc.nci.nih.gov/.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabs-remail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov  
Plate: LLM9254 row: f column: 02  
High quality sequence stop: 705.  
Location/Qualifiers

FEATURES  
source  
1. 792  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone\_image="IMAGE:4012801"  
/clone\_lib="NCI\_CGAP\_Lu30"  
/tissue\_type="tumor, metastatic to mammary"  
/lab\_host="DH10B"  
/note="Organ: lung; Vector: pCMV-SPORT6; Site: 1: NotI; Site: 2: SalI; transgenic model WNT-1, expression driven by MMTV-LTR enhancer; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies.  
Investigator providing samples: Gilbert Smith, NIH"

BASE COUNT 193 a 204 c 205 g 190 t

Query Match 20.0%; Score 35; DB 12; Length 792;  
Best Local Similarity 100.0%; Pred. No. 1.5e-07;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 TCTGGCAGCCAGCGCCCTTCTCTCAGCATGAGA 111  
|||||  
Db 659 TCTGGCAGCCAGCGCCCTTCTCTCAGCATGAGA 693

RESULT 6  
BB220676 307 bp mRNA linear EST 01-JUL-2000  
LOCUS BB220676 RIKEN full-length enriched, adult male aorta and vein Mus  
DEFINITION musculus cDNA clone A530061F23 3' similar to AF126063 Mus musculus connective tissue growth factor-like protein precursor (Ctgfl) mRNA  
mRNA sequence.

ACCESSION BB220676  
VERSION BB220676.1 GI:8889288  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus.

REFERENCE  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 307)

**AUTHORS**  
 Konno, H., Alzawa, K., Akahira, S., Akiyama, J., Arakawa, T., Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hirozane, T., Horii, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M., Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N., Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurihara, C., Kusakabe, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y., Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K., Shibata, K., Shibata, Y., Shigemoto, Y., Shitagawa, A., Shiraki, T., Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Tagawa, A., Takahashi, F., Tomioka, N., Toya, T., Tsunoda, Y., Watabiki, A., Watanabe, S., Yamamura, T., Yamane, T., Yano, R., Yasunishi, A., Yokota, T., Yoshida, K., Yoshiki, A., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.  
 RIKEN Mouse ESIS (Konno, H., et al.)  
 Unpublished (2000)  
**TITLE**  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
**JOURNAL**  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suenho-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-res@gsr.riken.go.jp,  
 URL: http://genome.gsc.riken.go.jp/  
**COMMENT**  
 Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoaka, S., Sasaki, N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 Thermostabilization and thermoactivation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)  
 Itoh, M., Katsunai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J., Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.  
 Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)  
 Carninci, P. and Hayashizaki, Y.  
 High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)  
 Please visit our web site (http://genome.rtc.riken.go.jp) for further details.  
**FEATURES**  
 source  
 Location/Qualifiers  
 1. 307  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /clone="A530061F23"  
 /clone\_id="RIKEN full-length enriched, adult male aorta and vein"  
 /sex="male"  
 /tissue\_type="aorta and vein"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /note="Site 1: SalI; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', GAGAGAGAGATCCAGAGCTCTTTTCTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 459.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGATCCAGAGCTCTTTTCTTTTNN 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pbluescript KS(+) after bulk excision from lambda phage I."

**Query Match**  
 Best Local Similarity 12.6%; Score 22; DB 10; Length 307;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

**BASE COUNT**  
 87 a 71 c 70 g 79 t  
**ORIGIN**

77 TCTGCGAGCGCGGCTTTCT 98  
 ||||||||||||||||||||  
 199 TCTGCGAGCGCGGCTTTCT 220

**RESULT 7**  
 LOCUS BF903057/c 299 bp mRNA linear EST 18-JAN-2001  
 DEFINITION CM3-MT0193-151200-565-e02 MT0193 Homo sapiens cDNA, mRNA sequence.  
 ACCESSION BF903057  
 VERSION BF903057.1 GI:12294516  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
**REFERENCE**  
 1 (bases 1 to 299)  
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bata, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.  
 Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
**TITLE**  
 Laboratory of Cancer Genetics  
**JOURNAL**  
 Ludwig Institute for Cancer Research  
**COMMENT**  
 Contact: Simpson A.J.G.  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM3&ct=CM3-MT0193-151200-565-e02&ct3=2000-12-15&ct4=1)  
 Seq primer: puc 18 forward  
 High quality sequence start: 8  
 High quality sequence stop: 299.  
**FEATURES**  
 source  
 Location/Qualifiers  
 1. 299  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_id="MT0193"  
 /dev\_stage="Adult"  
 /note="Organ: marrow; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196 716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

**Query Match**  
 Best Local Similarity 11.4%; Score 20; DB 12; Length 299;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

**BASE COUNT**  
 79 a 46 c 55 g 119 t  
**ORIGIN**

24 TTTTCAGAGATGCACATCTC 43  
 ||||||||||||||||||||  
 80 TTTTCAGAGATGCACATCTC 61

**RESULT 8**  
 LOCUS BB367824 313 bp mRNA linear EST 12-JUL-2000  
 DEFINITION BB367824 RIKEN full-length enriched, 16 days embryo head Mus musculus cDNA clone C130039K03 3, similar to Afl126053 Mus musculus connective tissue growth factor-like protein precursor (Ctgfl) mRNA





chemically-competent *E. coli* XL10-Gold (Stratagene) cells  
and selected for ampicillin resistance."

BASE COUNT 162 a 107 c 92 g 177 t  
ORIGIN

Query Match 11.4%; Score 20; DB 17; Length 538;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 94 TTCTCTTCAGCAGAGAAA 113  
|||||  
Db 310 TTCTCTTCAGCAGAGAAA 329

RESULT 10  
AG078128 683 bp DNA linear GSS 03-NOV-2001  
DEFINITION Pan troglodytes DNA, clone: PTB-072N12.R, genomic survey sequence.  
ACCESSION AG078128  
VERSION AG078128.1 GI:16629930  
KEYWORDS  
SOURCE Pan troglodytes male lymphoblast DNA, clone\_lib:PTB Chimpanzee Male  
BAC library clone:PTB-072N12.R.

## ORGANISM

REFERENCE 1  
AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,  
Tokoki, Y., Watanabe, H. and Sakaki, Y.  
TITLE BAC end sequences of library PTB  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 683)  
AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,  
Tokoki, Y., Watanabe, H. and Sakaki, Y.

TITLE Direct Submission  
JOURNAL Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical  
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
(E-mail: shirombes@sc.riken.go.jp, URL: http://bgp.gsc.riken.go.jp/,  
Tel: 81-45-503-9111, Fax: 81-45-503-9170)  
COMMENT Clones are derived from the chimpanzee BAC library PTB This BAC end  
was generated during the Rad process and may have higher chance of  
clone tracking errors.  
PRIMERS

Sequencing: MJ3Rev  
LIBRARY Vector : pKSI45  
R.Site 1 : SacI  
R.Site 2 : SacI

FEATURES  
source 1..683  
Location/Qualifiers  
/organism="Pan troglodytes"  
/db\_xref="taxon:9598"  
/clone="PTB-072N12.R"  
/sex="male"  
/cell\_type="lymphoblast"  
/clone\_lib="PTB Chimpanzee Male BAC library"  
BASE COUNT 198 a 65 c 187 g 226 t 7 others  
ORIGIN

Query Match 11.4%; Score 20; DB 17; Length 683;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 57 AACACGAGAGCTCCACAC 76  
|||||  
Db 657 AACACGAGAGCTCCACAC 676

RESULT 11  
AM660658 546 bp mRNA linear EST 25-APR-2001  
LOCUS  
DEFINITION 99634 MARC 1BOV Bos taurus cDNA 5', mRNA sequence.

ACCESSION AM660658  
VERSION AM660658.1 GI:7426485  
KEYWORDS  
SOURCE cow.  
ORGANISM Bos taurus

REFERENCE 1 (bases 1 to 546)  
AUTHORS Smith, R.P.L., Grosse, W.M., Freking, B.A., Roberts, A.J., Stone, R.T.,  
Casas, E., Wray, J.E., White, J., Cho, J., Fahrénkrug, S.C., Bennett,  
G.L., Heaton, M.P., Laegreid, W.W., Rohrer, G.A., Chitko-McKown, C.G.,  
Pettea, G., Holt, I., Karamycheva, S., Liang, F., Quackenbush, J. and  
Keefe, J.W.

TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA  
libraries and construction of a gene index for cattle  
JOURNAL Genome Res. 11 (4), 626-630 (2001)  
MEDLINE 21180013  
COMMENT Contact: Smith TPL  
USDA, ARS, US Meat Animal Research Center  
PO Box 166, Clay Center, NE 68933-0166, USA  
Tel: 402 762 4356  
Fax: 402 762 4390

Email: smith@mail.marc.usda.gov  
Single pass sequencing. Bases called and alt. trimmed with phred  
v0.980904.e. Vector identified by cross\_match with the -minscore 18  
and -mismatch 12 options.

PCR primers  
FORWARD: AGGAACAGCTATGACAT  
BACKWARD: GTTTCACAGTCAGCAG  
Plate: 90 row: L column: 23  
Seq primer: ATTAGTGACATATAG.  
FEATURES  
source 1..546  
Location/Qualifiers  
/organism="Bos taurus"  
/db\_xref="taxon:9913"  
/clone\_lib="MARC 1BOV"  
/tissue\_type="pooled"  
/lab\_host="DH10B"  
/note="Vector: pCMV SPORT6; Site\_1: NotI; Site\_2: SalI;  
Library made from pooled tissue from lymph node, ovary,  
fat, hypothalamus, and pituitary."

BASE COUNT 160 a 135 c 147 g 104 t  
ORIGIN

Query Match 10.9%; Score 19; DB 10; Length 546;  
Best Local Similarity 100.0%; Pred. No. 37;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 120 GGACAGCAGACTACTCTCC 138  
|||||  
Db 205 GGACAGCAGACTACTCTCC 223

RESULT 12  
BH112293/c 591 bp DNA linear GSS 19-JUN-2001  
LOCUS BH112293  
DEFINITION RPCI-24-265L18.TV RPCI-24 Mus musculus genomic clone RPCI-24-265L18  
, DNA sequence.  
ACCESSION BH112293  
VERSION BH112293.1 GI:14948266  
KEYWORDS GSS.  
SOURCE house mouse.  
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 591)  
AUTHORS Zhao, S., Nierman, W., Malek, J., Shatsman, S., Akinret, B., Levins, M.,  
Tsegaye, G., Geer, K., Krol, M., Shvartsbeyn, A., Gebregziorgis, E.,  
Russell, D., de Jong, P. and Fraser, C.M.  
Mouse BAC End Sequences from Library RPCI-24  
Unpublished (1999)  
other\_GSSs: RPCI-24-265L18.TVB

TITLE  
JOURNAL  
COMMENT

Contact: Shaying Zhao  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208

Email: szhao@tigr.org  
Clones are derived from the mouse BAC library RPCI-24. For BAC library availability, please contact Pletier de Jong (pdejong@tigr.org). Clones may be purchased from BACPAC Resources (<http://www.chori.org/bacpac/orderingframe.htm>). BAC end page: [http://www.tigr.org/tdb/bac\\_ends/mouse/bac\\_end\\_intro.html](http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html)  
Plate: 265 row: L column: 18  
Seq primer: T7  
Class: BAC ends.

# FEATURES

source  
Location/Qualifiers  
1..591  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone\_lib="RPCI-24-265L18"  
/clone\_1lb="RPCI-24"  
/sex="Male"  
/cell\_type="Spleen/Brain"  
/note="Vector: pTRABAC1; Site\_1: BamHI; Site\_2: BamHI; RPCI-24 Mouse BAC Library produced by Pletier de Jong. The library was cloned in the pTRABAC1 cloning vector at the BamHI sites using MboI partially digested male C57BL/6J DNA."  
BASE COUNT 210 a 114 c 121 g 146 t  
ORIGIN

Query Match 10.9%; Score 19; DB 17; Length 591;  
Best Local Similarity 100.0%; Pred. No. 38;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 63 GAAGGCTCCACACCTCTGG 81  
Db 59 GAAGGCTCCACACCTCTGG 41

RESULT 13  
BE778127/c 932 bp mRNA linear EST 20-OCT-2000  
LOCUS BE778127 601463356f1 NIH\_MGC.67 Homo sapiens cDNA clone IMAGE:3866748 5',  
DEFINITION mRNA sequence.  
ACCESSION BE778127  
VERSION BE778127.1 GI:10199325  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 932)  
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
TITLE NIH-MGC <http://mgc.nci.nih.gov/>.  
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)  
COMMENT Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Plate: L1AM9611 row: p column: 13  
High quality sequence stop: 633.  
Location/Qualifiers  
1..932

# FEATURES

source  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3866748"

/clone\_1lb="NIH\_MGC\_67"  
/issue\_type="retinoblastoma"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: eye; Vector: pCMV-SPORT6; Site\_1: NotI; Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.75 kb. Library constructed by Life Technologies."  
BASE COUNT 184 a 197 c 258 g 293 t  
ORIGIN

Query Match 10.9%; Score 19; DB 12; Length 932;  
Best Local Similarity 100.0%; Pred. No. 45;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 57 AAACAGAGGCTCCACAC 75  
Db 859 AAACAGAGGCTCCACAC 841

RESULT 14  
BI334722/c 957 bp mRNA linear EST 30-JUL-2001  
LOCUS BI334722 602999615f1 NIH\_MGC.12 Homo sapiens cDNA clone IMAGE:5141931 5',  
DEFINITION mRNA sequence.  
ACCESSION BI334722  
VERSION BI334722.1 GI:15019379  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 957)  
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
TITLE NIH-MGC <http://mgc.nci.nih.gov/>.  
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)  
COMMENT Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: Incyte Genomics, Inc.  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Plate: L1AM1349 row: m column: 04  
High quality sequence start: 145  
High quality sequence stop: 179.  
Location/Qualifiers  
1..957

# FEATURES

source  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5141931"  
/clone\_1lb="NIH\_MGC\_12"  
/issue\_type="cervical carcinoma cell line"  
/lab\_host="DH10B"  
/note="Organ: cervix; Vector: pCMV-SPORT6; Site\_1: NotI; Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.4 kb. Library prepared by Life Technologies."  
BASE COUNT 252 a 262 c 193 g 249 t 1 others  
ORIGIN

Query Match 10.9%; Score 19; DB 13; Length 957;  
Best Local Similarity 100.0%; Pred. No. 45;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 63 GAAGGCTCCACACCTCTGG 81.  
Db 322 GAAGGCTCCACACCTCTGG 304

RESULT 15  
CNS06MOV





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 15:19:16 ; Search time 1764.67 Seconds

(without alignments)  
10472.371 Million cell updates/sec

Title: US-10-010-408-1\_COPY\_1\_635

Perfect score: 635  
Sequence: 1 GACGCTCTGATCTCAGAG.....GTGATGACGTGCTTACC 635

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 2054640 seqs, 14551402878 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl:\*

1: gb\_ba:\*

2: gb\_bt:\*

3: gb\_in:\*

4: gb\_om:\*

5: gb\_ov:\*

6: gb\_pat:\*

7: gb\_ph:\*

8: gb\_pl:\*

9: gb\_pr:\*

10: gb\_ro:\*

11: gb\_sts:\*

12: gb\_sy:\*

13: gb\_un:\*

14: gb\_vl:\*

15: em\_ba:\*

16: em\_fun:\*

17: em\_hum:\*

18: em\_in:\*

19: em\_mu:\*

20: em\_om:\*

21: em\_or:\*

22: em\_ov:\*

23: em\_pat:\*

24: em\_ph:\*

25: em\_pl:\*

26: em\_ro:\*

27: em\_sts:\*

28: em\_un:\*

29: em\_vl:\*

30: em\_htg\_hum:\*

31: em\_htg\_inv:\*

32: em\_htg\_other:\*

33: em\_htg\_mus:\*

34: em\_htg\_dln:\*

35: em\_htg\_fod:\*

36: em\_htg\_mam:\*

37: em\_htg\_vrt:\*

38: em\_sy:\*

39: em\_htgo\_hum:\*

40: em\_htgo\_mus:\*

41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description        |
|------------|-------|-------------|--------|-------|--------------------|
| 1          | 218   | 34.3        | 137964 | 2     | AC126895 Rattus no |
| 2          | 218   | 34.3        | 226303 | 2     | AC095418 Rattus no |
| 3          | 166   | 26.1        | 137964 | 2     | AC126895 Rattus no |
| 4          | 114   | 18.0        | 1741   | 10    | AF259981 Rattus no |
| 5          | 90    | 14.2        | 1734   | 6     | AR210324 Rattus no |
| 6          | 90    | 14.2        | 1734   | 6     | AR210325 Rattus no |
| 7          | 90    | 14.2        | 1734   | 10    | AF100778 Rattus no |
| 8          | 90    | 14.2        | 61072  | 10    | AL731698 Rattus no |
| 9          | 90    | 14.2        | 216757 | 2     | AL669906 Mus muscu |
| 10         | 80    | 12.6        | 1739   | 10    | AF126063 Mus muscu |
| 11         | 32    | 5.0         | 738    | 6     | AR210337 Rattus no |
| 12         | 32    | 5.0         | 1266   | 6     | AX076919 Rattus no |
| 13         | 32    | 5.0         | 1266   | 6     | AX464186 Rattus no |
| 14         | 32    | 5.0         | 1283   | 6     | AF083500 Rattus no |
| 15         | 32    | 5.0         | 1293   | 6     | AR210322 Rattus no |
| 16         | 32    | 5.0         | 1293   | 6     | AR210323 Rattus no |
| 17         | 32    | 5.0         | 1309   | 9     | AF074604 Rattus no |
| 18         | 32    | 5.0         | 1427   | 9     | AF100780 Rattus no |
| 19         | 32    | 5.0         | 1450   | 9     | BC017782 Homo sapi |
| 20         | 32    | 5.0         | 107260 | 9     | AL139352 Human DNA |
| 21         | 27    | 4.3         | 51     | 6     | AR210371 Rattus no |
| 22         | 27    | 4.3         | 51     | 6     | AX076923 Rattus no |
| 23         | 22    | 3.5         | 142142 | 2     | AC091173 Rattus no |
| 24         | 22    | 3.5         | 149483 | 2     | AC110904 Rattus no |
| 25         | 22    | 3.5         | 186676 | 9     | AC067881 Rattus no |
| 26         | 22    | 3.5         | 200054 | 1     | AL646068 Rattus no |
| 27         | 21    | 3.3         | 150944 | 2     | AC112856 Rattus no |
| 28         | 21    | 3.3         | 154881 | 2     | AC122102 Rattus no |
| 29         | 21    | 3.3         | 181343 | 10    | AL671882 Rattus no |
| 30         | 21    | 3.3         | 215105 | 2     | AC073717 Mus muscu |
| 31         | 20    | 3.1         | 1003   | 5     | CHKTCEAA Rattus no |
| 32         | 20    | 3.1         | 1200   | 6     | AX122384 Rattus no |
| 33         | 20    | 3.1         | 1481   | 8     | AB033535 Rattus no |
| 34         | 20    | 3.1         | 1712   | 10    | AB008516 Rattus no |
| 35         | 20    | 3.1         | 36296  | 2     | HSE81G9 Rattus no  |
| 36         | 20    | 3.1         | 83021  | 2     | AC095541 Rattus no |
| 37         | 20    | 3.1         | 84001  | 9     | AL162739 Rattus no |
| 38         | 20    | 3.1         | 117840 | 9     | AL162400 Rattus no |
| 39         | 20    | 3.1         | 129090 | 2     | AC083961 Rattus no |
| 40         | 20    | 3.1         | 135817 | 2     | AC102328 Rattus no |
| 41         | 20    | 3.1         | 147990 | 2     | AC026144 Rattus no |
| 42         | 20    | 3.1         | 162486 | 2     | AC044779 Rattus no |
| 43         | 20    | 3.1         | 163258 | 2     | AP003514 Rattus no |
| 44         | 20    | 3.1         | 168050 | 9     | AC090526 Rattus no |
| 45         | 20    | 3.1         | 168566 | 2     | AC106167 Rattus no |

## ALIGNMENTS

RESULT 1  
AC126895 137964 bp DNA linear HTG 24-JUL-2002  
LOCUS Rattus norvegicus clone CH230-301E4, \*\*\* SEQUENCING IN PROGRESS  
DEFINITION \*\*\* 49 unordered pieces.  
ACCESSION AC126895  
VERSION AC126895.1 GI:21724040  
KEYWORDS HTG; HTGS; PHASE1.  
SOURCE Rattus norvegicus  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
REFERENCE 1 (bases 1 to 137964)  
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,

Alsbrooks, S.L., Amaralunge, H.C., Are, J.R., Ayele, M., Banks, T.,  
Barber, J., Benton, J., Bimage, K., Blankenburg, K., Brant, D.,  
Bouch, J., Bowle, S., Brileva, M., Brown, E., Brown, M., Bryant, N.P.,  
Buhay, C., Burck, P., Burkett, C., Burrell, R.E., Byrd, N.C.,  
Carion, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,  
Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,  
Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,  
Davila, N.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,  
Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,  
Doutlaw, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,  
Edwards, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,  
Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,  
Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,  
Correll, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,  
Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,  
Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B.,  
Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,  
Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,  
Karlisson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,  
Kralovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,  
Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Louised, H.,  
Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,  
Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,  
Massey, E., Mathiney, E., McLeod, M.P., Meador, M., Mel, G., Metzger, M.,  
Miner, G., Miner, Z., Mitchell, T., Mohabac, K., Morgan, M., Morris, S.,  
Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,  
Nguyen, N., Nickerson, E., Nwokwenkwo, S., Ogih, M., Okunolu, G.,  
Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,  
Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,  
Rives, M., Rojas, A., Rojibokan, T., Rolfe, M., Ruiz, S., Severy, G.,  
Scheer, S., Scott, G., Shen, H., Shooshari, N., Sisson, I.,  
Sodergren, E., Sonaik, T., Sparks, A., Stanley, H., Stone, H.,  
Sutton, A., Svatek, A., Taber, P., Tamerisa, A., Tamerisa, K., Tang, H.,  
Tansy, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,  
Usmani, K., Vasquez, L., Vera, V., Villalob, B., Vinson, R., Wang, O.,  
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,  
Williams, G., Williamson, A., Wleczek, R., Woden, S., Worley, K.,  
Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,  
Weinstock, G., and Gibbs, R.

Direct Submission  
Unpublished  
2 (bases 1 to 137964)  
Worley, K.C.

Direct Submission  
Submitted (10-JUL-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 137964)  
Worley, K.C.

Direct Submission  
Submitted (24-JUL-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: GZHG  
Center clone name: CH230-301E4  
----- Summary Statistics  
Sequencing vector: Plasmid:  
Chemistry: Dye-terminator Big Dye, 100% of reads  
Assembly program: Phrap, version 0.990329  
Consensus quality: 89474 bases at least Q40  
Consensus quality: 93422 bases at least Q30  
Consensus quality: 96506 bases at least Q20

----- NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank/draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 49 contigs. The true order of the pieces

\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 1056: contig of 1056 bp in length  
1057 1156: gap of unknown length  
1157 2335: contig of 1179 bp in length  
2336 2436: gap of unknown length  
2436 3442: contig of 1007 bp in length  
3442 3542: gap of unknown length  
3542 5081: contig of 1359 bp in length  
5081 5181: gap of unknown length  
5181 6424: contig of 1243 bp in length  
6424 6525: gap of unknown length  
6525 7815: contig of 1290 bp in length  
7815 7915: gap of unknown length  
7915 9153: contig of 1239 bp in length  
9153 9254: gap of unknown length  
9254 10445: contig of 1192 bp in length  
10445 10546: gap of unknown length  
10546 11772: contig of 1226 bp in length  
11772 11871: gap of unknown length  
11871 13240: contig of 1369 bp in length  
13240 13340: gap of unknown length  
13340 14359: contig of 1019 bp in length  
14359 14459: gap of unknown length  
14459 15651: contig of 1192 bp in length  
15651 15751: gap of unknown length  
15751 17494: contig of 1743 bp in length  
17494 17594: gap of unknown length  
17594 18679: contig of 1085 bp in length  
18679 20681: gap of unknown length  
20681 20781: contig of 1992 bp in length  
20781 22118: gap of unknown length  
22118 22218: gap of unknown length  
22218 23578: contig of 1360 bp in length  
23578 23678: gap of unknown length  
23678 25423: contig of 1745 bp in length  
25423 25523: gap of unknown length  
25523 27808: contig of 2285 bp in length  
27808 30272: gap of unknown length  
30272 30372: contig of 2364 bp in length  
30372 33091: gap of unknown length  
33091 33191: contig of 2719 bp in length  
33191 35777: gap of unknown length  
35777 35877: contig of 2586 bp in length  
35877 36908: gap of unknown length  
36908 37008: contig of 1031 bp in length  
37008 38930: gap of unknown length  
38930 39030: contig of 1922 bp in length  
39030 41210: gap of unknown length  
41210 41310: contig of 2180 bp in length  
41310 42937: gap of unknown length  
42937 46120: contig of 1627 bp in length  
46120 46220: gap of unknown length  
46220 48217: contig of 3083 bp in length  
48217 48317: gap of unknown length  
48317 50433: contig of 1997 bp in length  
50433 50533: gap of unknown length  
50533 52880: contig of 2116 bp in length  
52880 52980: gap of unknown length  
52980 55195: contig of 2347 bp in length  
55195 55294: gap of unknown length  
55294 57482: contig of 2214 bp in length  
57482 57582: gap of unknown length  
57582 60781: contig of 2188 bp in length  
60781 60881: gap of unknown length  
60881 62599: contig of 3199 bp in length  
62599 62599: contig of 1718 bp in length

```

* 62600 62699: gap of unknown length
* 62700 65203: contig of 2504 bp in length
* 65204 65304: gap of unknown length
* 65305 65914: contig of 1611 bp in length
* 65915 67015: gap of unknown length
* 67016 70839: contig of 3825 bp in length
* 70840 75139: gap of unknown length
* 75140 75239: contig of 4200 bp in length
* 75240 78297: gap of unknown length
* 78298 83138: contig of 4742 bp in length
* 83139 83239: gap of unknown length
* 83240 88205: contig of 4966 bp in length
* 88206 92238: gap of unknown length
* 92239 92338: contig of 3934 bp in length
* 92339 97339: gap of unknown length
* 97340 97439: gap of unknown length
* 103335 103634: contig of 6095 bp in length
* 103635 107080: gap of unknown length
* 107081 107180: gap of unknown length
* 107181 115202: contig of 8021 bp in length
* 115203 121237: gap of unknown length
* 121238 121337: contig of 5936 bp in length
* 121338 127910: gap of unknown length
* 127911 128010: contig of 6573 bp in length
* 128011 137964: gap of unknown length
* 137965 137966: contig of 9954 bp in length.
Location/Qualifiers
1. 137964
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/clone="CH230-301B4"

```

```

BASE COUNT 31831 a 34027 c 35229 g 31979 t 4898 others
ORIGIN

```

```

Query Match 34.3% Score 218; DB 2; Length 137964;
Best Local Similarity 100.0%; Pred. No. 3.2e-106;
Matches 218; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 308 GGTGTGTCCTGACGTCGTGCGACACCTGTACCTGTCTGTGACACACCCCACTGCCC 367
DB 104664 GGTGTGTCCTGACGTCGTGCGACACCTGTACCTGTCTGTGACACACCCCACTGCCC 104723
OY 368 ACAGGGGTCACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 427
DB 104724 ACAGGGGTCACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 104783
OY 428 GGGGAGTCCTGCGACACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 487
DB 104784 GGGGAGTCCTGCGACACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 104843
OY 488 TGGGGAGTCCTGCGACACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 525
DB 104844 TGGGGAGTCCTGCGACACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 104881

```

```

RESULT 2 AC095418 226303 bp DNA linear HTG 11-JUL-2002
LOCUS AC095418 Rattus norvegicus clone CH230-7C10, *** SEQUENCING IN PROGRESS ***
DEFINITION 51 unordered pieces.
ACCESSION AC095418
VERSION AC095418.3 GI:21717893
KEYWORDS HTG; HTGS_PHASE1.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 226303)

```

## AUTHORS

Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Albrooks,S.L., Amaralunge,H.C., Are,J.R., Ayele,M., Banks,T., Barbarta,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Bivava,M., Brown,E., Brown,M., Bryan,N.P., Bunah,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhury,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,T., Dim,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escoto,M., Falls,T., Ferraguto,D., Edwards,N., Ford,J., Foster,P., Frantz,P., Galsli,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hughes,M., Holloway,C., Hollins,B., Homsl,E., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Jarlssohn,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louised,H., Lozad,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapa,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mel,G., Metzger,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokwoko,S., Ogun,W., Okunodu,G., Oregunye,N., Oviado,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rivers,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savary,G., Scherer,S., Scott,G., Shen,H., Shoostari,N., Sisson,I., Sodergren,E., Sonake,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Syatek,A., Tabor,P., Tameis,A., Tameis,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Tellord,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinsom,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleceyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D., Weinstein,G. and Gibbs,R.

## Direct Submission

Unpublished  
2 (bases 1 to 226303)

## Direct Submission

Submitted (17-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

Submitted (11-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 226303)  
Worley,K.C.

## Direct Submission

Submitted (11-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On Jul 10, 2002 this sequence version replaced gi.17941885.

## Genome Center

Center: Baylor College of Medicine

## Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)

Center project name: GCDP

Center clone name: CH230-7C10

## Summary Statistics

Sequencing vector: Plasmid;

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 185950 bases at least Q40

Consensus quality: 190362 bases at least Q30

Consensus quality: 193076 bases at least Q20

\* NOTE: Estimated insert size may differ from sequence length  
(see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).

NOTE: This is a 'working draft' sequence. It currently consists of 51 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```

1 1051: contig of 1051 bp in length
1052 1151: gap of unknown length
1152 2286: contig of 1135 bp in length
2287 2386: gap of unknown length
2387 3433: contig of 1047 bp in length
3434 3533: gap of unknown length
3534 4614: contig of 1081 bp in length
4615 4714: gap of unknown length
4715 6581: contig of 1867 bp in length
6582 7773: gap of unknown length
7774 9530: contig of 1657 bp in length
9531 10651: contig of 1031 bp in length
10652 10761: gap of unknown length
10762 11958: contig of 1197 bp in length
11959 12058: gap of unknown length
12059 13140: contig of 1082 bp in length
13141 13240: gap of unknown length
13241 14777: contig of 1537 bp in length
14778 14877: gap of unknown length
14878 16063: contig of 1186 bp in length
16064 16163: gap of unknown length
16164 17585: contig of 1422 bp in length
17586 17685: gap of unknown length
17686 18916: contig of 1231 bp in length
18917 19016: gap of unknown length
19017 20305: contig of 1289 bp in length
20306 20405: gap of unknown length
20406 21537: contig of 1132 bp in length
21538 21637: gap of unknown length
21638 23703: contig of 2066 bp in length
23704 23803: gap of unknown length
23804 26306: contig of 2503 bp in length
26307 26406: gap of unknown length
26407 28431: contig of 2025 bp in length
28432 30615: contig of 2084 bp in length
30616 30715: gap of unknown length
30716 33215: contig of 2500 bp in length
33216 33315: gap of unknown length
33316 35357: contig of 2042 bp in length
35358 35457: gap of unknown length
35458 38012: contig of 2555 bp in length
38013 38112: gap of unknown length
38113 40338: contig of 2226 bp in length
40339 40438: gap of unknown length
40439 44051: contig of 3613 bp in length
44052 44151: gap of unknown length
44152 46294: contig of 2143 bp in length
46295 46395: gap of unknown length
46396 48845: contig of 2451 bp in length
48846 48945: gap of unknown length
48946 51753: contig of 2808 bp in length
51754 51853: gap of unknown length
51854 55915: contig of 4062 bp in length
55916 56015: gap of unknown length
56016 60468: contig of 4453 bp in length
60469 60568: gap of unknown length
60569 63640: contig of 3072 bp in length
63641 67888: contig of 4148 bp in length
67889 73014: gap of unknown length
73014: contig of 5026 bp in length

```

```

* 73015 73114: gap of unknown length
* 73115 73115 77313: contig of 4199 bp in length
* 77314 77313: gap of unknown length
* 77414 83786: contig of 6373 bp in length
* 83787 83886: gap of unknown length
* 83887 89590: contig of 5704 bp in length
* 89591 89690: gap of unknown length
* 89691 93946: contig of 4156 bp in length
* 93947 93946: gap of unknown length
* 93947 98126: contig of 4180 bp in length
* 98127 98226: gap of unknown length
* 98227 104258: contig of 6032 bp in length
* 104259 104358: gap of unknown length
* 104359 110994: contig of 6636 bp in length
* 110995 111094: gap of unknown length
* 111095 119838: contig of 8744 bp in length
* 119839 119938: gap of unknown length
* 119939 126339: contig of 6401 bp in length
* 126340 126439: gap of unknown length
* 126440 134888: contig of 8449 bp in length
* 134889 134988: gap of unknown length
* 134989 142709: contig of 7721 bp in length
* 142710 142809: gap of unknown length
* 142810 149818: contig of 7009 bp in length
* 149819 149918: gap of unknown length
* 149919 158719: contig of 8801 bp in length
* 158720 158819: gap of unknown length
* 158820 169688: contig of 11049 bp in length
* 169689 169968: gap of unknown length
* 169969 180390: contig of 10422 bp in length
* 180391 180490: gap of unknown length
* 180491 191272: contig of 10782 bp in length
* 191273 191372: gap of unknown length
* 191373 209667: contig of 18295 bp in length
* 209668 209767: gap of unknown length
* 209768 226303: contig of 16536 bp in length.

```

FEATURES  
source 1. .226303  
Location/Qualifiers

BASE COUNT 57875 a 53353 c 52917 g 57041 t 5117 others

Query Match 34.3%; Score 218; DB 2; Length 226303;

Best Local Similarity 100.0%; Pred. No. 2.9e-106;

Matches 218; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 308 GGTGTGTGCCCGCTGTGCGGACACCCGTACCTGTCTTGACACACCCAGTGGCC 367

DB 147645 GGTGTGTGCCCGCTGTGCGGACACCCGTACCTGTCTTGACACACCCAGTGGCC 147704

QY 368 ACAGGGGAGTACCCCTGTGCTGTGATGAGTGTGCTGTAAAGTGTGACGAGAGCT 427

DB 147705 ACAGGGGAGTACCCCTGTGCTGTGATGAGTGTGCTGTAAAGTGTGACGAGAGCT 147764

QY 428 GGGGAGTCTGCGACACCTGATGTGTGCGACCCACGAGGCGCTGTTCAGCC 487

DB 147765 GGGGAGTCTGCGACACCTGATGTGTGCGACCCACGAGGCGCTGTTCAGCC 147824

QY 488 TGGGGAGGCGCTTGGGCGGACATGGGCGTGTGTCTCT 525

DB 147825 TGGGGAGGCGCTTGGGCGGACATGGGCGTGTGTCTCT 147862

RESULT 3

AC126895 137964 bp DNA linear HTG 24-JUL-2002

DEFINITION Rattus norvegicus clone CH230-301E4, \*\*\* SEQUENCING IN PROGRESS

AC126895

ACCESSION AC126895

VERSION AC126895.1 GI:21724040

KEYWORDS HTG; HTGS\_PHASE1.

SOURCE Rattus norvegicus.



## ORGANISM

Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE  
AUTHORS

1 (bases 1 to 137964)

Muzny,D.M., Adams,C., Adio-Oduola,B., Alt-osman,F.R., Allen,C.,  
Albarracks,S.L., Amaralunge,H.C., Are,J.R., Ayele,M., Banks,T.,  
Barbarta,J., Benton,J., Blinige,K., Blankenburg,K., Bonini,D.,  
Bouck,J., Bowe,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,  
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,  
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,  
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,  
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,  
DeJana,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,  
Delaney,K.R., Delgado,O., Denn,A.B., Ding,Y., Dinh,H.H.,  
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,  
Earnharte,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,  
Falls,T., Ferraguto,D., Flagg,N., Ford,U., Foster,P., Frantz,P.,  
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,  
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,  
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,  
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,  
Honsi,F., Howard,S., Huber,J., Hulik,S., Hume,J., Jackson,L.E.,  
Jacobson,B., Jia,Y., Johnson,R., Jollivet,S., Joudah,S.,  
Karlssohn,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,  
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,  
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,  
Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,  
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,  
Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mel,G., Metzger,M.,  
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,  
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,N., Nguyen,N.,  
Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okunolu,G.,  
Orgunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,  
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,  
Rives,M., Rojas,A., Rojoubokan,I., Rolfe,M., Ruiz,S., Savary,G.,  
Scheiner,S., Scott,G., Shen,H., Shoshitari,N., Sisson,I.,  
Sodergren,E., Sonaik,T., Sparks,A., Stanley,H., Stone,H.,  
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,  
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,  
Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,  
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,  
Williams,G., Williamson,A., Wleczek,R., Wooden,S., Worley,K.,  
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
Weinstock,G. and Gidbs,R.

## TITLE

Direct Submission

## JOURNAL

Unpublished

## AUTHORS

2 (bases 1 to 137964)

## TITLE

Direct Submission

## JOURNAL

Submitted (10-JUN-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA

## REFERENCE

3 (bases 1 to 137964)

## AUTHORS

Worley,K.C.

## TITLE

Direct Submission

## JOURNAL

Submitted (24-JUN-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA

## COMMENT

Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

Project Information

Center project name: GZHG

Center clone name: CH230-301E4

Summary Statistics

Sequencing vector: Plasmid

Chemistry: Dye-terminator Big Dye: 100% of reads

Assembly program: Phrap, version 0.990329

Consensus quality: 89474 bases at least Q40

Consensus quality: 93422 bases at least Q30

Consensus quality: 96506 bases at least Q20

\* NOTE: Estimated insert size may differ from sequence length  
(see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_drift\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_drift_data.html)).  
\* NOTE: This is a "working draft" sequence. It currently  
\* consists of 49 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

|   |       |                             |
|---|-------|-----------------------------|
| 1 | 1056: | contig of 1056 bp in length |
| * | 1157  | gap of unknown length       |
| * | 2335  | contig of 1179 bp in length |
| * | 2435  | gap of unknown length       |
| * | 2436  | contig of 1007 bp in length |
| * | 3443  | gap of unknown length       |
| * | 3543  | contig of 1539 bp in length |
| * | 5082  | gap of unknown length       |
| * | 5182  | contig of 1243 bp in length |
| * | 6424  | gap of unknown length       |
| * | 6425  | contig of 1290 bp in length |
| * | 6525  | gap of unknown length       |
| * | 7815  | contig of 1239 bp in length |
| * | 7915  | gap of unknown length       |
| * | 9154  | contig of 1192 bp in length |
| * | 9254  | gap of unknown length       |
| * | 10445 | contig of 1192 bp in length |
| * | 10446 | gap of unknown length       |
| * | 10546 | contig of 1226 bp in length |
| * | 11771 | gap of unknown length       |
| * | 11872 | contig of 1369 bp in length |
| * | 13340 | gap of unknown length       |
| * | 13341 | contig of 1019 bp in length |
| * | 14359 | gap of unknown length       |
| * | 14360 | contig of 1192 bp in length |
| * | 14459 | gap of unknown length       |
| * | 1460  | contig of 1192 bp in length |
| * | 15652 | gap of unknown length       |
| * | 15752 | contig of 1743 bp in length |
| * | 17495 | gap of unknown length       |
| * | 17595 | contig of 1085 bp in length |
| * | 18679 | gap of unknown length       |
| * | 18780 | contig of 1902 bp in length |
| * | 20682 | gap of unknown length       |
| * | 20782 | contig of 1337 bp in length |
| * | 22119 | gap of unknown length       |
| * | 22218 | contig of 1360 bp in length |
| * | 23219 | gap of unknown length       |
| * | 23578 | contig of 1360 bp in length |
| * | 23579 | gap of unknown length       |
| * | 23679 | contig of 1745 bp in length |
| * | 25423 | gap of unknown length       |
| * | 25424 | contig of 2285 bp in length |
| * | 27808 | gap of unknown length       |
| * | 27908 | contig of 2364 bp in length |
| * | 27909 | gap of unknown length       |
| * | 30272 | contig of 2364 bp in length |
| * | 30273 | gap of unknown length       |
| * | 30373 | contig of 2719 bp in length |
| * | 33091 | gap of unknown length       |
| * | 33191 | contig of 2566 bp in length |
| * | 33577 | gap of unknown length       |
| * | 35778 | contig of 1031 bp in length |
| * | 35877 | gap of unknown length       |
| * | 36908 | contig of 1922 bp in length |
| * | 36909 | gap of unknown length       |
| * | 38930 | contig of 1922 bp in length |
| * | 39030 | gap of unknown length       |
| * | 39031 | contig of 2180 bp in length |
| * | 41210 | gap of unknown length       |
| * | 41211 | contig of 1627 bp in length |
| * | 41310 | gap of unknown length       |
| * | 41311 | contig of 1627 bp in length |
| * | 42937 | gap of unknown length       |
| * | 43037 | contig of 3063 bp in length |
| * | 43038 | gap of unknown length       |
| * | 46120 | contig of 1997 bp in length |
| * | 46220 | gap of unknown length       |
| * | 48217 | contig of 1997 bp in length |
| * | 48317 | gap of unknown length       |
| * | 50433 | contig of 2116 bp in length |
| * | 50533 | gap of unknown length       |
| * | 50534 | contig of 2347 bp in length |
| * | 52880 | gap of unknown length       |
| * | 52980 | contig of 2214 bp in length |
| * | 52981 | gap of unknown length       |

\* 55195 55294: gap of unknown length  
\* 55295 57482: contig of 2188 bp in length  
\* 57483 57582: gap of unknown length  
\* 57583 60781: contig of 3199 bp in length  
\* 60782 60881: gap of unknown length  
\* 60882 62599: contig of 1718 bp in length  
\* 62600 62700: gap of unknown length  
\* 62700 65203: contig of 2504 bp in length  
\* 65203 65303: gap of unknown length  
\* 65304 66914: contig of 1611 bp in length  
\* 66915 67015: gap of unknown length  
\* 67015 70839: contig of 3825 bp in length  
\* 70840 70939: gap of unknown length  
\* 70940 75139: contig of 4200 bp in length  
\* 75140 75239: gap of unknown length  
\* 75240 78296: contig of 3057 bp in length  
\* 78297 78396: gap of unknown length  
\* 78397 83138: contig of 4742 bp in length  
\* 83139 83238: gap of unknown length  
\* 83239 88204: contig of 4966 bp in length  
\* 88205 88304: gap of unknown length  
\* 88305 92238: contig of 3934 bp in length  
\* 92239 92339: gap of unknown length  
\* 92339 97339: contig of 5001 bp in length  
\* 97340 97439: gap of unknown length  
\* 97440 103534: contig of 6095 bp in length  
\* 103535 103634: gap of unknown length  
\* 103635 107080: contig of 3446 bp in length  
\* 107081 107180: gap of unknown length  
\* 107181 115201: contig of 8021 bp in length  
\* 115202 115301: gap of unknown length  
\* 115302 121237: contig of 5936 bp in length  
\* 121238 121337: gap of unknown length  
\* 121338 127910: contig of 6573 bp in length  
\* 127911 137964: gap of unknown length  
\* 137964 137964: contig of 9954 bp in length.  
Location/Qualifiers  
1. 137964  
/organism="Rattus norvegicus"  
/db\_xref="taxon:10116"  
/clone="CH230-301E4"

BASE COUNT 31831 a 34027 c 35229 g 31979 t 4898 others  
ORIGIN

Query Match 26.1%; Score 166; DB 2; Length 137964;  
Best Local Similarity 99.5%; Pred. No. 3.1e-78;  
Matches 216; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGGCTTGTGATCCAGAGACCCCTGGGGTGGAGACGGGCTTGSCAAGCTGCAGCC 60  
Db 74321 GAGGCTTGTGATCCAGAGACCCCTGGGGTGGAGACGGGCTTGSCAAGCTGCAGCC 74262  
QY 61 GCTGGGAGTGGCTTGAATGAGAGCTTATTACTGGGAAGTGAAGAGAGCTC 120  
Db 74261 GCTGGGAGTGGCTTGAATGAGAGCTTATTACTGGGAAGTGAAGAGAGCTC 74202  
QY 121 CTGTCAAGCTTGTCTAAAGTCTTACGACCTTGTGGGCTTGSCAAGCTGCAGA 180  
Db 74201 CTGTCAAGCTTGTCTAAAGTCTTACGACCTTGTGGGCTTGSCAAGCTGCAGA 74142  
QY 181 CACCTTGTGGTGGGCTGCACGGCTCACCCTCAGGT 217  
Db 74141 CACCTTGTGGTGGGCTGCACGGCTCACCCTCAGGT 74105

RESULT 4  
AF259981 1741 bp mRNA linear ROD 09-MAY-2000  
LOCUS AF259981  
DEFINITION Rattus norvegicus CCN family protein COP-1 (COP-1) mRNA, complete  
ACCESSION AF259981  
VERSION AF259981.1 GI:7739780  
KEYWORDS

SOURCE Rattus norvegicus.  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
REFERENCE 1 (bases 1 to 1741)  
Zhang, R., Averdukh, L., Zhu, W., Zhang, H., Jo, H., Dempsey, P. J.,  
Coffey, R. J., Pardue, A. B., and Liang, P.  
Identification of rCOP-1, a new member of the CCN protein family,  
as a negative regulator for cell transformation  
Mol. Cell. Biol. 18 (10), 6131-6141 (1998)  
JOURNAL 98414629  
MEDLINE 9742130  
PUBMED 9742130  
REFERENCE 2 (bases 1 to 1741)  
Liang, P.  
TITLE Direct Submission  
JOURNAL Submitted (24-APR-2000) Cell Biology, Vanderbilt-Ingram Cancer  
Center, 649 MRB II, Nashville, TN 37232, USA  
FEATURES  
source  
1. 1741  
Location/Qualifiers  
/organism="Rattus norvegicus"  
/db\_xref="taxon:10116"  
1. 1741  
/gene="COP-1"  
262..1014  
/gene="COP-1"  
/note="secreted protein"  
/codon\_start=1  
/product="CCN family protein COP-1"  
/protein\_id="AF69011.1"  
/db\_xref="GI:7739781"  
/translation="MRGSPILRLATSFLLSMVCAQLCRPCTGPPPPGCGV  
LVLDGGCKKVCARLRLTESCEHLHVCPESSGLVCGPAGGGAVALDDEDDGEV  
NGRRYLDGEFFKPCRVLCSCDDGFTCLPLCSEDTLPSPMDPRPKRIQVPGCCPE  
WVCDGVTPAIRSAAGHOLSAIVTPASADAPMPWMSRAMGCSSTGIGITATRVSN  
QNRFCLEIQRRLCLPRPCLAARSHSMNSAF"

BASE COUNT 386 a 491 c 480 g 384 t  
ORIGIN

Query Match 18.0%; Score 114; DB 10; Length 1741;  
Best Local Similarity 99.1%; Pred. No. 7e-50;  
Matches 214; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 213 CAGGTTGAAGCTGTGCTCCAGAGGACGAGATGAGGGGACGCCATGATCCAT 272  
Db 226 CAGGTTGAAGCTGTGCTCCAGAGGACGAGATGAGGGGACGCCATGATCCAT 285  
QY 273 CTTCGTGGCAGCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 332  
Db 286 CTTCGTGGCAGCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 345  
QY 333 CCCTGTACCTGTGCTTCCTTGACACCCAGTGGCCACAGAGGGGTACCCCTGTGTGAT 392  
Db 346 CCCTGTACCTGTGCTTCCTTGACACCCAGTGGCCACAGAGGGGTACCCCTGTGTGAT 405  
QY 393 GGCTGTGGCTGTGTAAAGTGTGTGACAGGAGGCTG 428  
Db 406 GGCTGTGGCTGTGTAAAGTGTGTGACAGGAGGCTG 441

RESULT 5  
AR210324 1734 bp DNA linear PAT 20-JUN-2002  
LOCUS AR210324  
DEFINITION Sequence 17 from patent US 6387657.  
ACCESSION AR210324  
VERSION AR210324.1 GI:21512525  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1734)  
Botstein, D. A., Cohen, R. L., Goddard, A. D., Gurney, A. L., Hillan, K. J.,  
Lawrence, D. A., Levine, A. J., Pennica, D., Roy, M. Ann. and Wood, W. I.

TITLE WISP polypeptides and nucleic acids encoding same  
JOURNAL Patent: US 6387657-A 17 14-May-2002;  
FEATURES Location/Qualifiers  
source 1..1734  
BASE COUNT 355 a 491 c 495 g 393 t  
ORIGIN

Query Match 14.2%; Score 90; DB 6; Length 1734;  
Best Local Similarity 100.0%; Pred. No. 5.8e-37;  
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 410 AGTGTGTGACGAGAGCTGGGGAGTCTCGACACCACTGCATGCTGGAGCCAGCCCA 469  
|||||  
DB 418 AGTGTGTGACGAGAGCTGGGGAGTCTCGACACCACTGCATGCTGGAGCCAGCCCA 477  
|||||

QY 470 GGGCCTGGTTGTGACCTGGGGGAGGCC 499  
|||||  
DB 478 GGGCCTGGTTGTGACCTGGGGGAGGCC 507  
|||||

RESULT 6  
AR210325/c 1734 bp DNA linear PAT 20-JUN-2002  
LOCUS AR210325 18 from patent US 6387657.  
DEFINITION AR210325  
ACCESSION AR210325  
VERSION AR210325.1 GI:21512526  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1734)  
AUTHORS Botstein,D.A., Cohen,R.L., Goddard,A.D., Gurney,A.L., Hillan,K.J.,  
Lawrence,D.A., Levine,A.J., Pennica,D., Roy,M.Ann. and Wood,W.I.  
TITLE WISP polypeptides and nucleic acids encoding same  
JOURNAL Patent: US 6387657-A 18 14-May-2002;  
FEATURES Location/Qualifiers  
source 1..1734  
BASE COUNT 393 a 495 c 491 g 355 t  
ORIGIN

Query Match 14.2%; Score 90; DB 6; Length 1734;  
Best Local Similarity 100.0%; Pred. No. 5.8e-37;  
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 410 AGTGTGTGACGAGAGCTGGGGAGTCTCGACACCACTGCATGCTGGAGCCAGCCCA 469  
|||||  
DB 418 AGTGTGTGACGAGAGCTGGGGAGTCTCGACACCACTGCATGCTGGAGCCAGCCCA 477  
|||||

QY 470 GGGCCTGGTTGTGACCTGGGGGAGGCC 499  
|||||  
DB 1257 GGGCCTGGTTGTGACCTGGGGGAGGCC 1228  
|||||

RESULT 7  
AF100778 1734 bp mRNA linear ROD 17-DEC-1998  
LOCUS AF100778  
DEFINITION Mus musculus connective tissue growth factor related protein WISP-2  
(WISP2) mRNA, complete cds.  
ACCESSION AF100778  
VERSION AF100778.1 GI:4028578  
KEYWORDS  
SOURCE Mus musculus.  
ORGANISM Mus musculus.  
REFERENCE 1 (bases 1 to 1734)  
AUTHORS Lee,J., Brush,J., Taneyhill,L.A., Deuel,B., Lew,M., Wetanabe,C.,  
Cohen,R.L., Melhem,M.F., Finley,G.G., Quirke,P., Goddard,A.D.,  
Hillan,K.J., Gurney,A.L., Botstein,D. and Levine,A.J.  
TITLE WISP genes are members of the connective tissue growth factor

JOURNAL MEDLINE  
PUBMED 9843955  
REFERENCE 2 (bases 1 to 1734)  
AUTHORS Pennica,D.  
TITLE Direct Submission  
JOURNAL Submitted (23-OCT-1998) Molecular Oncology, Genentech, Inc., 1 DNA  
Way, South San Francisco, CA 94080, USA  
Location/Qualifiers  
source 1..1734  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/cell\_line="C57MG"  
/cell\_type="epithelial"  
/tissue\_type="mammary"  
/note="Transformed by Wnt-1"  
1..1734  
/gene="Wisp2"  
257..1012  
/gene="Wisp2"  
/codon\_start=1  
/product="connective tissue growth factor related protein  
WISP-2"  
/protein\_id="AAC96320.1"  
/db\_xref="GI:4028579"  
/translation="MRGNPLHLHLAISFLCISWYSQLCRAPCPPTPPQCPGVP  
LVLDGCGCCRVACARRLGSCDHLHVCDSQGLVCPGAGPSRGAVCLFEDDSCEV  
NGRRYIDETFRPNCRLVLCRDGDGFTCLPLCSBVRALPSMCPRPRIIOVGRCCPE  
WYCDVAVQPAIOPSSAOGHOLSVTPASADGPPNMSTAMGPCSTTGILATRVIS  
NMRFCOEIDRIICLSRPLASRHSWNSAF"

BASE COUNT 355 a 491 c 495 g 393 t  
ORIGIN

Query Match 14.2%; Score 90; DB 10; Length 1734;  
Best Local Similarity 100.0%; Pred. No. 5.8e-37;  
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 410 AGTGTGTGACGAGAGCTGGGGAGTCTCGACACCACTGCATGCTGGAGCCAGCCCA 469  
|||||  
DB 418 AGTGTGTGACGAGAGCTGGGGAGTCTCGACACCACTGCATGCTGGAGCCAGCCCA 477  
|||||

QY 470 GGGCCTGGTTGTGACCTGGGGGAGGCC 499  
|||||  
DB 478 GGGCCTGGTTGTGACCTGGGGGAGGCC 507  
|||||

RESULT 8  
AL731698 61072 bp DNA linear ROD 24-MAY-2002  
LOCUS AL731698  
DEFINITION Mouse DNA sequence from clone Rp23-161B3 on chromosome 2, complete  
sequence.  
ACCESSION AL731698  
VERSION AL731698.10 GI:21214309  
KEYWORDS HTG.  
SOURCE house mouse.  
ORGANISM Mus musculus.  
REFERENCE 1 (bases 1 to 61072)  
AUTHORS Wallis,J.  
TITLE Direct Submission  
JOURNAL Submitted (24-MAY-2002) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
humquery@sanger.ac.uk Clone requests: clonequest@sanger.ac.uk  
On May 25, 2002 this sequence version replaced gi:21213601.  
During sequence assembly data is compared from overlapping clones.  
Where differences are found these are annotated as variations  
together with a note of the overlapping clone name. Note that the  
variation annotation may not be found in the sequence submission  
corresponding to the overlapping clone, as we submit sequences with  
only a small overlap as described above.





SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 738)  
AUTHORS Botstein,D.A., Cohen,R.L., Goddard,A.D., Gurney,A.L., Hillan,K.J., Lawrence,D.A., Levine,A.J., Pernice,D., Roy,M.Ann. and Wood,W.I., WISP polypeptides and nucleic acids encoding same  
JOURNAL Patent: US 6387657-A 38 14-MAY-2002;  
FEATURES Location/Qualifiers  
source 1..738  
BASE COUNT 104 a 272 c 238 g 124 t  
ORIGIN

Query Match 5.0%; Score 32; DB 6; Length 738;  
Best Local Similarity 100.0%; Pred. No. 1.1e-05;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 375 GTACCCCTGCTGCTGATGGCTGCTGCTG 406  
DB 115 GTACCCCTGCTGCTGATGGCTGCTGCTG 146

RESULT 12  
AX076919 1266 bp DNA linear PAT 22-FEB-2001  
LOCUS Sequence 31 from Patent WO0105836.  
DEFINITION AX076919  
ACCESSION AX076919  
VERSION AX076919.1 GI:13121575  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
AUTHORS Botstein,D., Goddard,A., Gurney,A.L., Hillan,K.J., Roy,M.A. and Wood,W.I.  
TITLE Polypeptidic compositions and methods for the treatment of tumors  
JOURNAL Patent: WO 0105836-A 31 25-JAN-2001;  
FEATURES Location/Qualifiers  
source 1..1266  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 216 a 418 c 390 g 242 t  
ORIGIN

Query Match 5.0%; Score 32; DB 6; Length 1266;  
Best Local Similarity 100.0%; Pred. No. 1e-05;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 375 GTACCCCTGCTGCTGATGGCTGCTGCTG 406  
DB 136 GTACCCCTGCTGCTGATGGCTGCTGCTG 167

RESULT 13  
AX464186 1266 bp DNA linear PAT 16-JUL-2002  
LOCUS Sequence 319 from Patent WO0140466.  
DEFINITION AX464186  
ACCESSION AX464186  
VERSION AX464186.1 GI:21899109  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
AUTHORS Baker,K.P., Beresini,M., Deforge,L., Desjoviers,L., Filvaroff,E., Gao,M.Q., Gerlitsen,M.E., Goddard,A., Godowski,P.J., Gurney,A.L., Sherwood,S., Smith,V., Stewart,T.A., Tums,D., Watanabe,C.K., Wood,W.L. and Zhang,Z.  
TITLE Secreted and transmembrane polypeptides and nucleic acids encoding

JOURNAL same  
Patent: WO 0140466-A 319 07-JUN-2001;  
Genentech Inc. (US)  
FEATURES Location/Qualifiers  
source 1..1266  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 216 a 418 c 390 g 242 t  
ORIGIN

Query Match 5.0%; Score 32; DB 6; Length 1266;  
Best Local Similarity 100.0%; Pred. No. 1e-05;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 375 GTACCCCTGCTGCTGATGGCTGCTGCTG 406  
DB 136 GTACCCCTGCTGCTGATGGCTGCTGCTG 167

RESULT 14  
AF083500 1283 bp mRNA linear PRI 04-NOV-1998  
LOCUS AF083500  
DEFINITION Homo sapiens connective tissue growth factor-like protein  
ACCESSION AF083500  
VERSION AF083500.1 GI:3462835  
KEYWORDS  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
AUTHORS Kumar,S., Hand,A.T., Connor,J.C., Dadds,R.A., Ryan,P.J., Trill,J.G., Fisher,S.M., Slemmon,J.R., Lipshutz,D.B., Bartholomew,V., James,I.E., Rieman,D.J., Gowen,M. and Lee,J.C.  
TITLE Identification and cloning of CTGF-L from human osteoblasts, a novel cysteine rich protein containing an IGF binding domain  
JOURNAL Bone 23 (5), S240 (1998)  
REFERENCE Kumar,S.  
AUTHORS Direct Submission  
TITLE Submitted (11-AUG-1998) Bone & Cartilage Biology, UM 2109, SmithKline Beecham, 709 Swedeland Road, King of Prussia, PA 19406, USA  
JOURNAL

FEATURES Location/Qualifiers  
source 1..1283  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="20"  
/map="20q12-q13"  
/cell\_type="primary osteoblast"  
9..761  
/note="CTGF-L: encodes IGF binding (IGFB), Von Willebrand factor type C (VWC) and thrombospondin type I (TSP1) domains; member of the CCN (CTGF/cyrl6/nov) family; lacks the fourth carboxy-terminal domain present in other members of the CCN family"  
/codon\_start=1  
/product="connective tissue growth factor-like protein precursor"  
/protein\_id="AAC70350.1"  
/db\_xref="GI:3462836"  
/translation="MRGTPKTHLAFSLCLSKVRYTQLCPPTCTCPMPPRPCPLGVP LVLDGCGCVCARLGEPCDQDQHVCDASQGLVCGAGAGGALCLLAEDSSCEV NGRLYREGTEPQHSIRRCRDGDFCTCPLCSSEVRLPSMDCPPHRRREYVKGCCPE WYCGGGGIGTGPPLPAQGPQPSGLYSSLPVPGPCPEWSTANMPCSTTGGGLMATVSN QNRFRLEFRRLCLSRPCPPSRGSSPONSAP"  
BASE COUNT 235 a 418 c 389 g 241 t  
ORIGIN

Query Match 5.0%; Score 32; DB 9; Length 1283;  
Best Local Similarity 100.0%; Pred. No. 1e-05;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 375 GTACCCCTGGTGTGATGGCTGTGCTGTG 406  
 ||||||||||||||||||||||||||||  
 Db 135 GTACCCCTGGTGTGATGGCTGTGCTGTG 166

## RESULT 15

AR210322

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

Matches

32; Conservative

0; Mismatches

0; Indels

0; Gaps

0;

QY

375 GTACCCCTGGTGTGATGGCTGTGCTGTG 406

|||||

Db

148 GTACCCCTGGTGTGATGGCTGTGCTGTG 179

Search completed: July 28, 2003, 21:14:43  
 Job time : 1766.67 secs





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 14:31:11 ; Search time 166.775 Seconds  
(without alignments)  
8574.520 Million cell updates/sec

Title: US-10-010-408-1\_COPY\_1\_635

Perfect score: 635

Sequence: 1 GACGCTCTGATCTCCAGAG.....GTGATGACGTGCTTCACC 635

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 2185239 seqs, 1125999159 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

N.GeneSeq\_101002:\*

1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*

2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*

3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*

4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*

5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*

6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*

7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*

8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*

9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*

10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*

11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*

12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*

13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*

14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*

15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*

16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*

17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*

18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*

19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*

20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*

21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*

22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*

23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*

24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1          | 635   | 100.0       | 1708   | 20    | AAZ07516    |
| 2          | 387   | 60.9        | 753    | 20    | AAZ07517    |
| 3          | 318   | 50.1        | 681    | 20    | AAZ07521    |
| 4          | 210   | 33.1        | 210    | 20    | AAZ07519    |
| 5          | 90    | 14.2        | 177    | 20    | AAZ07518    |
| 6          | 90    | 14.2        | 753    | 20    | AAZ76489    |
| 7          | 90    | 14.2        | 1734   | 20    | AAZ76488    |
| 8          | 32    | 5.0         | 738    | 20    | AAZ76501    |
| 9          | 32    | 5.0         | 750    | 20    | AAZ76487    |

|    |    |     |       |    |          |
|----|----|-----|-------|----|----------|
| 10 | 32 | 5.0 | 1257  | 20 | AAZ28435 |
| 11 | 32 | 5.0 | 1266  | 21 | AAA30048 |
| 12 | 32 | 5.0 | 1266  | 22 | AAZ21403 |
| 13 | 32 | 5.0 | 1266  | 22 | AAZ60368 |
| 14 | 32 | 5.0 | 1266  | 22 | AAZ97451 |
| 15 | 32 | 5.0 | 1267  | 21 | AAZ77565 |
| 16 | 32 | 5.0 | 1285  | 19 | AAZ29260 |
| 17 | 32 | 5.0 | 1293  | 20 | AAZ76486 |
| 18 | 32 | 5.0 | 1309  | 22 | AAZ8214  |
| 19 | 32 | 5.0 | 1337  | 22 | AAZ6952  |
| 20 | 32 | 5.0 | 1352  | 22 | AAZ46936 |
| 21 | 32 | 5.0 | 1522  | 20 | AAZ16595 |
| 22 | 32 | 5.0 | 2136  | 22 | AAZ94706 |
| 23 | 32 | 5.0 | 13255 | 20 | AAZ76842 |
| 24 | 27 | 4.3 | 51    | 20 | AAZ76535 |
| 25 | 27 | 4.3 | 51    | 21 | AAZ77569 |
| 26 | 27 | 4.3 | 51    | 21 | AAZ30051 |
| 27 | 27 | 4.3 | 51    | 22 | AAZ60371 |
| 28 | 27 | 4.3 | 51    | 22 | AAZ97454 |
| 29 | 23 | 3.6 | 634   | 22 | AAZ91750 |
| 30 | 23 | 3.6 | 634   | 22 | AAZ93768 |
| 31 | 20 | 3.1 | 1200  | 22 | AAZ67265 |
| 32 | 20 | 3.1 | 34980 | 22 | AAZ68531 |
| 33 | 19 | 3.0 | 203   | 24 | ABN22295 |
| 34 | 19 | 3.0 | 217   | 24 | ABL74433 |
| 35 | 19 | 3.0 | 318   | 22 | AAZ26172 |
| 36 | 19 | 3.0 | 332   | 22 | AAZ18242 |
| 37 | 19 | 3.0 | 372   | 18 | AAZ63339 |
| 38 | 19 | 3.0 | 372   | 22 | AAZ59399 |
| 39 | 19 | 3.0 | 405   | 22 | AAZ99220 |
| 40 | 19 | 3.0 | 425   | 17 | AAZ03735 |
| 41 | 19 | 3.0 | 443   | 22 | AAZ08427 |
| 42 | 19 | 3.0 | 616   | 21 | AAZ16215 |
| 43 | 19 | 3.0 | 816   | 24 | AAZ38997 |
| 44 | 19 | 3.0 | 816   | 24 | ABK84404 |
| 45 | 19 | 3.0 | 816   | 24 | ABL68621 |

#### ALIGNMENTS

|          |  |                          |
|----------|--|--------------------------|
| RESULT 1 | AAZ07516   | standard; cDNA; 1708 BP. |
| ID       | AAZ07516   |                          |
| AC       | AAZ07516:  |                          |
| DT       | 26-NOV-1999  | (first entry)            |
| DE       | Rat HICP polypeptide encoding cDNA.  |                          |
| KW       | Heparin-induced CCN-1-like protein; HICP; cell-associated activity; ss; cardiovascular disorder; aberrant cell proliferation; fibrotic disorder. |                          |
| OS       | Rattus sp.   |                          |
| PN       | W09947556-A2.  |                          |
| PD       | 23-SEP-1999.   |                          |
| PF       | 18-MAR-1999;   | 99WO-US05999.            |
| PR       | 19-MAR-1998;   | 98US-0044273.            |
| PA       | (TUFTS ) TUFTS COLLEGE.  |                          |
| PI       | Castellot JJ;  |                          |
| PI       | WPI: 1999-562060/47.   |                          |
| DR       | P-PSDB; AAZ7434.   |                          |
| XX       | Nucleic acid sequences encoding rat heparin-induced CCN-1-like protein, used in methods to identify modulators or in diagnostic applications     |                          |
| PT       |  |                          |



AAZ07521  
ID AAZ07521 standard; cDNA; 681 BP.  
XX  
AC AAZ07521;  
XX  
DT 26-NOV-1999 (first entry)  
XX  
DE Rat HICP mature polypeptide coding sequence.  
XX  
KW Heparin-induced CCN-like protein; HICP; cell-associated activity; ss;  
KW cardiovascular disorder; aberrant cell proliferation; fibrotic disorder.  
XX  
OS Rattus sp.  
XX  
PN MO9947556-A2.  
XX  
PD 23-SEP-1999.  
XX  
PE 18-MAR-1999; 99WO-US05999.  
XX  
PR 19-MAR-1998; 98US-0044273.  
XX  
PA (TUFT ) TUFTS COLLEGE.  
XX  
PI Castellet JJ;  
XX  
DR WPI: 1999-562060/47.  
DR P-PSDB: AAY27440.  
XX  
PT Nucleic acid sequences encoding rat heparin-induced CCN-like protein,  
XX used in methods to identify modulators or in diagnostic applications  
XX  
PS Disclosure; Fig 2; 108pp; English.  
XX  
CC The invention provides a rat heparin-induced CCN-like protein (HICP)  
CC protein. Agents that stimulate or inhibit HICP protein activity or  
CC expression, antisense HICP nucleic acid molecules and HICP antibodies,  
CC can be used to modulate cell-associated activity. HICP modulators can be  
CC used to treat disorders characterized by aberrant HICP protein activity  
CC or expression. Probes capable of hybridizing to HICP mRNA or antibodies  
CC specific for HICP can be used to detect HICP activity in a biological  
CC sample. HICP can be used to treat disorders, such as a cardiovascular or  
CC fibrotic disorder, characterized by aberrant cell proliferation. The  
CC present sequence represents the coding sequence of rat HICP mature  
CC polypeptide.  
XX  
CC  
XX  
SO Sequence 681 BP; 121 A; 211 C; 215 G; 134 T; 0 other;  
Query Match 50.1%; Score 318; DB 20; Length 681;  
Best Local Similarity 100.0%; Pred. No. 1.3e-145;  
Matches 318; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 318 CAGCTGTGCGGACACCTGTACTGTCTTGACACACCCAGTCCACAGGGGTA 377  
DB 1 CAGCTGTGCGGACACCTGTACTGTCTTGACACACCCAGTCCACAGGGGTA 60  
QY 378 CCCCTGGTGTGATGGCTGTGGCTGCTGTAAGTGTGACGAGGCTGGGGAGTCC 437  
DB 61 CCCCTGGTGTGATGGCTGTGGCTGCTGTAAGTGTGACGAGGCTGGGGAGTCC 120  
QY 438 TGGACACACCTGCATGTCTGCGACCCAGCCAGGGGCTGTTGTCAAGCTGGGGCAGGC 497  
DB 121 TGGACACACCTGCATGTCTGCGACCCAGCCAGGGGCTGTTGTCAAGCTGGGGCAGGC 180  
QY 498 CCTGGCGGCATGGGGCTGTGTCTCTTGATGATGACGATGAGTGTGAGGTGAT 557  
DB 181 CCTGGCGGCATGGGGCTGTGTCTCTTGATGATGACGATGAGTGTGAGGTGAT 240  
QY 558 GGCGGACAGTCTGATGAGAGACCTTTAAACCAATTCAGAGGCTCTTGCCGCTGT 617  
DB 241 GGCGGACAGTCTGATGAGAGACCTTTAAACCAATTCAGAGGCTCTTGCCGCTGT 300  
QY 618 GATGACGGTGGCTTCACC 635

DB 301 GATGACGGTGGCTTCACC 318  
|||||  
RESULT 4  
AAZ07519  
ID AAZ07519 standard; cDNA; 210 BP.  
XX  
AC AAZ07519;  
XX  
DT 26-NOV-1999 (first entry)  
XX  
DE Rat HICP IGFBP domain encoding cDNA.  
XX  
KW Heparin-induced CCN-like protein; HICP; cell-associated activity; ss;  
KW cardiovascular disorder; aberrant cell proliferation; fibrotic disorder;  
KW insulin-like growth factor binding protein; IGFBP.  
XX  
OS Rattus sp.  
XX  
PN MO9947556-A2.  
XX  
PD 23-SEP-1999.  
XX  
PE 18-MAR-1999; 99WO-US05999.  
XX  
PR 19-MAR-1998; 98US-0044273.  
XX  
PA (TUFT ) TUFTS COLLEGE.  
XX  
PI Castellet JJ;  
XX  
DR WPI: 1999-562060/47.  
DR P-PSDB: AAY27438.  
XX  
PT Nucleic acid sequences encoding rat heparin-induced CCN-like protein,  
XX used in methods to identify modulators or in diagnostic applications  
XX  
PS Disclosure; Page 104; 108pp; English.  
XX  
CC The invention provides a rat heparin-induced CCN-like protein (HICP)  
CC protein. Agents that stimulate or inhibit HICP protein activity or  
CC expression, antisense HICP nucleic acid molecules and HICP antibodies,  
CC can be used to modulate cell-associated activity. HICP modulators can be  
CC used to treat disorders characterized by aberrant HICP protein activity  
CC or expression. Probes capable of hybridizing to HICP mRNA or antibodies  
CC specific for HICP can be used to detect HICP activity in a biological  
CC sample. HICP can be used to treat disorders, such as a cardiovascular or  
CC fibrotic disorder, characterized by aberrant cell proliferation. The  
CC present sequence represents a nucleotide sequence encoding the insulin-  
CC like growth factor binding protein (IGFBP) domain of HICP polypeptide.  
XX  
CC  
XX  
SO Sequence 210 BP; 27 A; 65 C; 74 G; 44 T; 0 other;  
Query Match 33.1%; Score 210; DB 20; Length 210;  
Best Local Similarity 100.0%; Pred. No. 9.9e-93;  
Matches 210; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 318 CAGCTGTGCGGACACCTGTACTGTCTTGACACACCCAGTCCACAGGGGTA 377  
DB 1 CAGCTGTGCGGACACCTGTACTGTCTTGACACACCCAGTCCACAGGGGTA 60  
QY 378 CCCCTGGTGTGATGGCTGTGGCTGCTGTAAGTGTGACGAGGCTGGGGAGTCC 437  
DB 61 CCCCTGGTGTGATGGCTGTGGCTGCTGTAAGTGTGACGAGGCTGGGGAGTCC 120  
QY 438 TGGACACACCTGCATGTCTGCGACCCAGCCAGGGGCTGTTGTCAAGCTGGGGCAGGC 497  
DB 121 TGGACACACCTGCATGTCTGCGACCCAGCCAGGGGCTGTTGTCAAGCTGGGGCAGGC 180  
QY 498 CCTGGCGGCATGGGGCTGTGTCTCTTG 527  
DB 181 CCTGGCGGCATGGGGCTGTGTCTCTTG 210

RESULT 5  
ID AA207518 standard; CDNA; 177 BP.  
XX  
AC AA207518;  
XX  
DT 26-NOV-1999 (first entry)  
XX  
DE Rat HICP VMC domain encoding CDNA.  
XX  
KM Hepatin-induced CCN-like protein; HICP, cell-associated activity; ss;  
KM Cardiovascular disorder; aberrant cell proliferation; fibrotic disorder;  
KM Von Willebrand C domain; VMC.  
XX  
OS Rattus sp.  
XX  
PN M09947556-A2.  
XX  
PD 23-SEP-1999.  
XX  
PF 18-MAR-1999; 99MO-US05999.  
XX  
PR 19-MAR-1998; 98US-0044273.  
XX  
PA (TUFT ) TUFTS COLLEGE.  
XX  
PI Castellet JJ;  
XX  
DR WPI: 1999-562060/47.  
DR P-PSDB; AAY27436.  
XX  
PT Nucleic acid sequences encoding rat heparin-induced CCN-like protein,  
XX used in methods to identify modulators or in diagnostic applications.  
XX  
PS Disclosure; Page 103; 108bp; English.  
XX  
XX The invention provides a rat heparin-induced CCN-like protein (HICP)  
CC protein. Agents that stimulate or inhibit HICP protein activity or  
CC expression, antisense HICP nucleic acid molecules and HICP antibodies,  
CC can be used to modulate cell-associated activity. HICP modulators can be  
CC used to treat disorders characterized by aberrant HICP protein activity  
CC or expression. Probes capable of hybridizing to HICP mRNA or antibodies  
CC specific for HICP can be used to detect HICP activity in a biological  
CC sample. HICP can be used to treat disorders, such as a cardiovascular or  
CC fibrotic disorder, characterized by aberrant cell proliferation. The  
CC present sequence represents a cDNA encoding the Von Willebrand C (VMC)  
CC a domain of the HICP polypeptide.  
XX  
SQ Sequence 177 BP; 35 A; 47 C; 60 G; 35 T; 0 other;  
XX  
Query Match 14.2%; Score 90; DB 20; Length 177;  
Best Local Similarity 100.0%; Pred. No. 5.5e-34;  
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
QY 546 TGTGAGTGAATGCCGACGACTGATGAGAGAGACCTTTAAACCAATTCAGAGTC 605  
DB 1 TGTGAGTGAATGCCGACGACTGATGAGAGAGACCTTTAAACCAATTCAGAGTC 60  
XX  
QY 606 CTGTGCCCTGTGATGACGCTGCCTCACC 635 1  
DB 61 CTGTGCCCTGTGATGACGCTGCCTCACC 90  
XX  
RESULT 6  
AA76489/C  
ID AAX76489 standard; DNA; 753 BP.  
XX  
AC AAX76489;  
XX  
DT 06-AUG-1999 (first entry)  
XX

DE Mouse WISP-2 protein complementary nucleotide sequence SEQ ID NO:18.  
XX  
XX WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;  
KM connective tissue growth factor; cancer; melanoma; arteriosclerosis;  
KM leukaemia; lymphoid malignancy; haematopoiesis-related disorder;  
KM tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;  
KM kidney disorder; bone-related disorder; osteoporosis; trauma; burn;  
KM connective tissue disorder; catabolic state; inflammation;  
KM testicular-related disorder; angiogenesis; immunological disorder; ss.  
XX  
OS Mus sp.  
XX  
PN M09921998-A1.  
XX  
PD 06-MAY-1999.  
XX  
PF 29-OCT-1998; 98MO-US22991.  
XX  
PR 14-APR-1998; 98US-0081695.  
PR 29-OCT-1997; 97US-0063704.  
PR 03-FEB-1998; 98US-0073612.  
XX  
PA (GETH ) GENENTECH INC.  
XX  
PI Botstein DA, Cohen RL, Goddard A, Gurney AL, Hillan K;  
PI Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WI;  
XX  
DR WPI: 1999-337420/28.  
XX  
XX New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3  
PT  
PS Disclosure; Page 179-180; 284pp; English.  
XX  
XX The present invention describes Wnt-1 induced secreted polypeptides,  
CC WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2  
CC and WISP-3 have homology to connective tissue growth factor (CTGF).  
CC Products from the present invention can be used to treat WISP-related  
CC disorders such as breast, ovarian, and colon cancer or melanoma. The  
CC products can be used to treat arteriosclerosis. The products can also be  
CC used to treat other diseases e.g. benign and malignant tumours,  
CC leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,  
CC hypothalamic and other glandular, macrophagal, epithelial, stromal, and  
CC blastocoele disorders, haematopoiesis-related disorders, tissue-growth  
CC disorders, skin disorders, desmoplasia, fibrotic lesions, kidney  
CC disorders, bone-related disorders such as osteoporosis, trauma such as  
CC burns, incisions, and other wounds, connective tissue disorders,  
CC catabolic states, testicular-related disorders, and inflammatory,  
CC angiogenic and immunologic disorders including arteriosclerosis. The  
CC products can also be used for detection and diagnosis especially of  
CC individuals with neoplastic cell growth or proliferation. The products  
CC can be used in the production of transgenic or knock-out animals.  
CC Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing  
CC cells.  
XX  
SQ Sequence 753 BP; 159 A; 229 C; 238 G; 127 T; 0 other;  
XX  
Query Match 14.2%; Score 90; DB 20; Length 753;  
Best Local Similarity 100.0%; Pred. No. 5.3e-34;  
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
QY 410 AGTGTGTGACGAGGCTGGGGAGTCTGCGACACACCTGCATGTTGGACCCAGCCCA 469  
DB 592 AGTGTGTGACGAGGCTGGGGAGTCTGCGACACACCTGCATGTTGGACCCAGCCCA 533  
XX  
QY 470 GGGCTGTGTTTGTGACGCTGGGGAGGCGCC 499  
DB 532 GGGCTGTGTTTGTGACGCTGGGGAGGCGCC 503  
XX  
RESULT 7  
AA76488  
ID AAX76488 standard; DNA; 1734 BP.  
XX

AC AAX76488;  
 XX  
 DT 06-AUG-1999 (first entry)  
 XX  
 DE Mouse WISP-2 protein nucleotide sequence SEQ ID NO:17.  
 XX  
 XX WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;  
 KW connective tissue growth factor; cancer; melanoma; arteriosclerosis;  
 KW leukaemia; lymphoid malignancy; hematopoiesis-related disorder;  
 KW tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;  
 KW kidney disorder; bone-related disorder; osteoporosis; trauma; burn;  
 KW connective tissue disorder; catabolic state; inflammation;  
 KW testicular-related disorder; angiogenesis; immunological disorder; ss.  
 XX  
 OS Mus sp.  
 XX  
 XX WO921998-A1.  
 XX  
 PD 06-MAY-1999.  
 XX  
 XX 29-OCT-1998; 98MO-US22991.  
 XX  
 XX 14-APR-1998; 98US-0081695.  
 PR 29-OCT-1997; 97US-0063704.  
 PR 03-FEB-1998; 98US-0073612.  
 XX  
 XX (GENENTECH INC.  
 XX Botstein DA, Cohen RL, Goddard A, Gurney AL, Hillan K;  
 PI Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WI;  
 XX  
 DR WPI; 1999-337420/28.  
 XX P-PSDB; AAV17651.  
 XX  
 PT New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3  
 XX  
 XX Example 2; Page 178-179; 284pp; English.  
 XX  
 XX The present invention describes Wnt-1 induced secreted polypeptides,  
 CC WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2  
 CC and WISP-3 have homology to connective tissue growth factor (CTGF).  
 CC Products from the present invention can be used to treat WISP-related  
 CC disorders such as breast, ovarian, and colon cancer or melanoma. The  
 CC products can be used to treat arteriosclerosis. The products can also be  
 CC used to treat other diseases e.g. benign and malignant tumours,  
 CC leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,  
 CC hypothalamic and other glandular, macrophagal, epithelial, stromal, and  
 CC blastocoeleic disorders, hematopiasis-related disorders, tissue-growth  
 CC disorders, skin disorders, desmoplasia, fibrotic lesions, kidney  
 CC disorders, bone-related disorders such as osteoporosis, trauma such as  
 CC burns, incisions, and other wounds, connective tissue disorders,  
 CC catabolic states, testicular-related disorders, and inflammatory,  
 CC angiogenic and immunologic disorders including arteriosclerosis. The  
 CC products can also be used for detection and diagnosis especially of  
 CC individuals with neoplastic cell growth or proliferation. The products  
 CC can be used in the production of transgenic or knock-out animals.  
 CC Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing  
 CC cells.  
 CC  
 XX  
 XX Sequence 1734 BP; 355 A; 491 C; 495 G; 393 T; 0 other;  
 SQ  
 Query Match 14.2%; Score 90; DB 20; Length 1734;  
 Best Local Similarity 100.0%; Pred. No. 5.2e-34;  
 Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 410 AGTGTGTCACGAGGAGTCTGTGACCACTGTGTCGAGCCAGCCCA 469  
 DB 418 AGTGTGTCACGAGGAGTCTGTGACCACTGTGTCGAGCCAGCCCA 477  
 OY 470 GGGCCTGTTGTTCAGCTGGGGGAGGCC 499  
 DB 478 GGGCCTGTTGTTCAGCTGGGGGAGGCC 507

RESULT 8  
 ID AAX76501 standard; DNA; 738 BP.  
 XX  
 XX AAX76501;  
 XX  
 DT 06-AUG-1999 (first entry)  
 XX  
 DE Human WISP-2 protein nucleotide sequence clone SEQ ID NO:38.  
 XX  
 XX WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;  
 KW connective tissue growth factor; cancer; melanoma; arteriosclerosis;  
 KW leukaemia; lymphoid malignancy; hematopoiesis-related disorder;  
 KW tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;  
 KW kidney disorder; bone-related disorder; osteoporosis; trauma; burn;  
 KW connective tissue disorder; catabolic state; inflammation;  
 KW testicular-related disorder; angiogenesis; immunological disorder; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO921998-A1.  
 XX  
 PD 06-MAY-1999.  
 XX  
 XX 29-OCT-1998; 98MO-US22991.  
 XX  
 XX 14-APR-1998; 98US-0081695.  
 PR 29-OCT-1997; 97US-0063704.  
 PR 03-FEB-1998; 98US-0073612.  
 XX  
 XX (GENENTECH INC.  
 XX Botstein DA, Cohen RL, Goddard A, Gurney AL, Hillan K;  
 PI Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WI;  
 XX  
 DR WPI; 1999-337420/28.  
 XX  
 PT New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3  
 XX  
 XX Example 4; Page 199-200; 284pp; English.  
 XX  
 XX The present invention describes Wnt-1 induced secreted polypeptides,  
 CC WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2  
 CC and WISP-3 have homology to connective tissue growth factor (CTGF).  
 CC Products from the present invention can be used to treat WISP-related  
 CC disorders such as breast, ovarian, and colon cancer or melanoma. The  
 CC products can be used to treat arteriosclerosis. The products can also be  
 CC used to treat other diseases e.g. benign and malignant tumours,  
 CC leukaemia and lymphoid malignancies, neuronal, glial, astrocytal,  
 CC hypothalamic and other glandular, macrophagal, epithelial, stromal, and  
 CC blastocoeleic disorders, hematopiasis-related disorders, tissue-growth  
 CC disorders, skin disorders, desmoplasia, fibrotic lesions, kidney  
 CC disorders, bone-related disorders such as osteoporosis, trauma such as  
 CC burns, incisions, and other wounds, connective tissue disorders,  
 CC catabolic states, testicular-related disorders, and inflammatory,  
 CC angiogenic and immunologic disorders including arteriosclerosis. The  
 CC products can also be used for detection and diagnosis especially of  
 CC individuals with neoplastic cell growth or proliferation. The products  
 CC can be used in the production of transgenic or knock-out animals.  
 CC Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing  
 CC cells.  
 CC  
 XX  
 XX Sequence 738 BP; 104 A; 272 C; 238 G; 124 T; 0 other;  
 SQ  
 Query Match 5.0%; Score 32; DB 20; Length 738;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-05;  
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 375 GTACCCCTGTGCTGTGATGCTGTGCTGTGCTG 406  
 DB 115 GTACCCCTGTGCTGTGATGCTGTGCTGTGCTG 146

RESULT 9  
AAx76487/c  
ID AAX76487 standard; DNA: 750 BP.  
XX  
AC AAX76487;  
XX  
DT 06-AUG-1999 (first entry)  
XX  
DE Human WISP-2 protein complementary nucleotide sequence SEQ ID NO:14.  
XX  
KM WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;  
KM connective tissue growth factor; cancer; melanoma; arteriosclerosis;  
KM leukemia; lymphoid malignancy; haematopoiesis-related disorder;  
KM tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;  
KM kidney disorder; bone-related disorder; osteoporosis; trauma; burn;  
KM connective tissue disorder; catabolic state; inflammation;  
KM testicular-related disorder; angiogenesis; immunological disorder; ss.  
XX  
OS Homo sapiens.  
XX  
PN W09921998-A1.  
XX  
PD 06-MAY-1999.  
XX  
PF 29-OCT-1998; 98WO-US22991.  
XX  
PR 14-APR-1998; 98US-0081695.  
PR 29-OCT-1997; 97US-0063704.  
PR 03-FEB-1998; 98US-0073612.  
XX  
PA (GETH ) GENENTECH INC.  
XX  
PI Botstein DA, Cohen RL, Goddard A, Gurney AL, Hillan K;  
PI Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WI;  
XX  
DR WPI: 1999-337420/28.  
XX  
PT New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3  
XX  
PS Disclosure; Page 175-176; 284pp; English.  
XX  
CC The present invention describes Wnt-1 induced secreted polypeptides,  
CC WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2  
CC and WISP-3 have homology to connective tissue growth factor (CTGF).  
CC Products from the present invention can be used to treat WISP-related  
CC disorders such as breast, ovarian, and colon cancer or melanoma. The  
CC products can be used to treat arteriosclerosis. The products can also be  
CC used to treat other diseases e.g. benign and malignant tumours,  
CC leukemia and lymphoid malignancies, neuronal, glial, astrocytal,  
CC hypochromic and other glandular, macrophagal, epithelial, stromal, and  
CC blastococcal disorders, haematopoiesis-related disorders, tissue-growth  
CC disorders, skin disorders, desmoplasia, fibrotic lesions, kidney  
CC disorders, bone-related disorders such as osteoporosis, trauma such as  
CC burns, incisions, and other wounds, connective tissue disorders,  
CC catabolic states, testicular-related disorders, and inflammatory,  
CC angiogenic and immunologic disorders including arteriosclerosis. The  
CC products can also be used for detection and diagnosis especially of  
CC individuals with neoplastic cell growth or proliferation. The products  
CC can be used in the production of transgenic or knock-out animals.  
CC Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing  
CC cells.  
XX  
SQ Sequence 750 BP; 125 A; 242 C; 274 G; 109 T; 0 other;

Query Match 5.0%; Score 32; DB 20; Length 750;  
Best Local Similarity 100.0%; Pred. No. 1.3e-05;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

375 GTACCCCTGCTGCTGATGCTGCTGCTG 406  
|||||  
624 GTACCCCTGCTGCTGATGCTGCTGCTG 593

RESULT 10  
AAX28435  
ID AAX28435 standard; DNA: 1257 BP.  
XX  
AC AAX28435;  
XX  
DT 22-JUN-1999 (first entry)  
XX  
DE EGF-like homologue PRO261 coding sequence.  
XX  
KM Antibody; PRO187; PRO533; PRO214; PRO240; PRO211; PRO230; PRO261; PRO246;  
KM EBAF-2; inhibitor; tumour growth; cancer; EGF-like homologue;  
KM FGF-8 homologue; ss.  
XX  
OS Homo sapiens.  
XX  
PN W09914327-A2.  
XX  
PD 25-MAR-1999.  
XX  
PF 10-SEP-1998; 98WO-US18824.  
XX  
PR 25-NOV-1997; 97US-0066840.  
PR 17-SEP-1997; 97US-0059114.  
PR 17-SEP-1997; 97US-0059117.  
PR 18-SEP-1997; 97US-0059263.  
PR 15-OCT-1997; 97US-0062125.  
PR 17-OCT-1997; 97US-0062285.  
PR 17-OCT-1997; 97US-0062287.  
PR 24-OCT-1997; 97US-0062816.  
PR 29-OCT-1997; 97US-0063704.  
XX  
PA (GETH ) GENENTECH INC.  
XX  
PI Botstein D, Goddard A, Gurney A, Hillan K, Lawrence DA;  
PI Roy M, Wood WI;  
XX  
DR WPI: 1999-229532/19.  
XX  
PT P-PSDB; AAY05285.  
XX  
PT Antibodies against specific proteins overexpressed in tumours  
XX  
PS Example 1; Fig 22; 130pp; English.  
XX  
CC This sequence encodes the EGF-like homologue PRO261.  
CC The invention relates to antibodies (Ab) that bind to any of the  
CC polypeptides (1)-designated PRO187, PRO533, PRO214, PRO240, PRO211;  
CC PRO230; PRO261; PRO246 or EBAF-2. The Ab, or other agents that inhibit  
CC expression and/or activity of (1) are used: (1) to inhibit growth of  
CC tumours; and (11) as diagnostic/prognostic reagents for detection or  
CC quantification of (1) in cells or tissues, by standard immunoassays, with  
CC overexpression being indicative of cancer. For therapeutic use, the Ab  
CC may be conjugated to a toxin, chemotherapeutic agent or radioisotope.  
CC Genes expressing (1), many of which are growth factor homologues, are  
CC overexpressed in some cases of cancer.  
XX  
SQ Sequence 1257 BP; 215 A; 416 C; 385 G; 241 T; 0 other;

Query Match 5.0%; Score 32; DB 20; Length 1257;  
Best Local Similarity 100.0%; Pred. No. 1.3e-05;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

375 GTACCCCTGCTGCTGATGCTGCTGCTG 406  
|||||  
127 GTACCCCTGCTGCTGATGCTGCTGCTG 158

RESULT 11  
AAA30048  
ID AAA30048 standard; cDNA: 1266 BP.  
XX  
AC AAA30048;

XX 09-AUG-2000 (first entry)  
 DT Human PRO261 nucleotide sequence.  
 XX  
 DE  
 XX  
 KW Antibody; PRO187; PRO533; PRO214; PRO240; PRO211; PRO230; PRO261; PRO246;  
 KW PRO317; tumour growth inhibitor; cancer; diagnosis; treatment; human;  
 KW cell growth; proliferation; growth factor; ADPEPT;  
 KW antibody dependent enzyme mediated prodrug therapy; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200015666-A2.  
 XX  
 PD 23-MAR-2000.  
 XX  
 PE 08-SEP-1999; 99WO-US20594.  
 XX  
 PR 10-SEP-1998; 98US-0099803.  
 PR 10-SEP-1998; 98WO-US18824.  
 XX  
 PA (GETH ) GENENTECH INC.  
 XX  
 PI Goddard A, Gurney AL, Hillan KJ, Roy MA, Wood WI, Bolstein D;  
 DR WPI: 2000-271386/23.  
 DR P-PSDB; AAY88573.  
 XX  
 PT New isolated antibodies which bind to specific polypeptides used for  
 PT diagnosis and treatment of neoplastic cell growth and proliferation -  
 XX  
 PS Example 7; Fig 13; 200pp; English.  
 XX  
 CC This sequence represents a human PRO261 nucleotide sequence. PRO261 is a  
 CC growth factor. The invention relates to isolated antibodies which bind to  
 CC a polypeptide. The "PRO" polypeptides are encoded by genes which are over  
 CC expressed in the genome of tumour cells. Vectors and host cells  
 CC comprising the nucleic acid encoding the antibodies are used in the  
 CC production of the antibodies. The antibodies and nucleic acids encoding  
 CC them are used for diagnosing a tumour in a mammal. The antibodies are  
 CC used for inhibiting the growth of tumour cells and identifying compounds  
 CC that inhibit a biological or immunological activity of and/or expression  
 CC of a PRO187, PRO533, PRO214, PRO240, PRO211, PRO230, PRO261, PRO246 or  
 CC PRO317 polypeptide. The antibody can be used in antibody dependent enzyme  
 CC mediated prodrug therapy (ADPEPT) by conjugating the antibody to a  
 CC prodrug-activating enzyme which converts a prodrug to an anti-cancer  
 CC drug. The antibodies can be fluorescently labelled and monitored by light  
 CC microscopy, flow cytometry or fluorimetry for diagnosis and prognosis of  
 CC tumours.  
 CC  
 CC Sequence 1266 BP; 216 A; 418 C; 390 G; 242 T; 0 other:  
 XX  
 SO  
 Query Match 5.0%; Score 32; DB 21; Length 1266;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-05;  
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 375 GTACCCCTGTCGTGATGCTGTGCTGCTG 406  
 ||||||||||||||||||||||||||||||||  
 Db 136 GTACCCCTGTCGTGATGCTGTGCTGCTG 167

KW cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;  
 KW adipocyte; A-peptide; factor VIIA; gene therapy; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200140466-A2.  
 XX  
 PD 07-JUN-2001.  
 XX  
 PE 01-DEC-2000; 2000WO-US32678.  
 XX  
 PR 01-DEC-1999; 99WO-US28301.  
 PR 01-DEC-1999; 99WO-US28634.  
 PR 02-DEC-1999; 99WO-US28551.  
 PR 02-DEC-1999; 99WO-US28564.  
 PR 02-DEC-1999; 99WO-US28565.  
 PR 09-DEC-1999; 99US-0170262.  
 PR 16-DEC-1999; 99WO-US30095.  
 PR 20-DEC-1999; 99WO-US30911.  
 PR 20-DEC-1999; 99WO-US30999.  
 PR 30-DEC-1999; 99WO-US31243.  
 PR 06-JAN-2000; 2000WO-US00277.  
 PR 06-JAN-2000; 2000WO-US00376.  
 PR 11-FEB-2000; 2000WO-US03565.  
 PR 18-FEB-2000; 2000WO-US04341.  
 PR 18-FEB-2000; 2000WO-US04342.  
 PR 22-FEB-2000; 2000WO-US04414.  
 PR 24-FEB-2000; 2000WO-US04914.  
 PR 24-FEB-2000; 2000WO-US05004.  
 PR 01-MAR-2000; 2000WO-US05601.  
 PR 20-MAR-2000; 2000WO-US07377.  
 PR 21-MAR-2000; 2000WO-US07532.  
 PR 30-MAR-2000; 2000WO-US08439.  
 PR 17-MAY-2000; 2000WO-US13705.  
 PR 22-MAY-2000; 2000WO-US14042.  
 PR 30-MAY-2000; 2000WO-US14941.  
 PR 02-JUN-2000; 2000WO-US15264.  
 PR 10-NOV-2000; 2000WO-US30873.  
 XX  
 PA (GETH ) GENENTECH INC.  
 XX  
 PI Baker KP, Beresini M, DeForge L, Desnoyers L, Filvaroff E, Gao W;  
 PI Gerlitsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;  
 PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;  
 DR WPI: 2001-408281/43.  
 DR P-PSDB; AAU12331.  
 XX  
 PT Isolated, secretory and transmembrane PRO polypeptide used to detect  
 PT other PRO polypeptides, link bioactive molecules to cells expressing  
 PT PRO polypeptides, and detect the presence of mammalian tumours e.g.  
 PT lung, breast, prostate, cervical -  
 XX  
 PS Claim 3; Fig 319; 813pp; English.  
 XX  
 CC AAS21244-AAS21518 encode for novel human secretory and transmembrane  
 CC PRO polypeptides. The PRO polypeptides are useful to detect other  
 CC PRO polypeptides, to link bioactive molecules to cells expressing  
 CC PRO polypeptides, and to detect the presence of mammalian lung, colon,  
 CC breast, prostate, rectal, cervical or liver tumours by comparing PRO  
 CC polypeptide expression in a cell sample to that in a control sample.  
 CC Some of the 275 sequences are also useful to stimulate the release of  
 CC tumour necrosis factor-alpha (TNF-alpha) from human blood, the  
 CC proliferation or differentiation of chondrocytes, the proliferation or  
 CC gene expression in pericyte cells, the release of proteoglycans from  
 CC cartilage, the proliferation of inner ear utricular supporting cells or  
 CC of T-lymphocytes, the release of a cytokine from peripheral blood  
 CC monocytes (PMCs), or the proliferation of endothelial cells. Some of  
 CC the PRO polypeptides may modulate glucose or free fatty acid uptake by  
 CC skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide  
 CC to factor VIIA. The PRO polypeptides can be used in assays to identify  
 CC molecules involved in binding interactions. The polynucleotides encoding

CC PRO polypeptides can be used to generate probes, antisense RNA/DNA,  
CC transgenic or knock out animals and can be used in gene therapy.

XX  
XX Sequence 1266 BP; 216 A; 418 C; 390 G; 242 T; 0 other;

Query Match 5.0%; Score 32; DB 22; Length 1266;

Best Local Similarity 100.0%; Pred. No. 1.3e-05;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 375 GTACCCCTGCTGCTGATGCTGCTGCTG 406  
Db 136 GTACCCCTGCTGCTGATGCTGCTGCTG 167

RESULT 13  
AA60368  
ID AAF60368 standard; cDNA; 1266 BP.  
XX  
XX AAF60368;  
AC  
XX 27-APR-2001 (first entry)  
DT  
XX  
XX PRO261 coding sequence.  
DE  
XX  
XX Cytostatic; PRO protein; tumour; cancer; ss.  
KM  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO200105836-A1.  
PN  
XX  
XX 25-JAN-2001.  
PD  
XX  
XX 20-DEC-1999; 99WO-US30999.  
PF  
XX  
XX 20-JUL-1999; 99US-0144758.  
PR 26-JUL-1999; 99US-0145698.  
XX  
XX 08-SEP-1999; 99WO-US20594.  
PR 13-SEP-1999; 99WO-US20944.  
XX  
XX 15-SEP-1999; 99WO-US21090.  
PR 05-OCT-1999; 99WO-US23089.  
XX  
XX 29-NOV-1999; 99WO-US28214.  
PR 30-NOV-1999; 99WO-US28313.  
XX  
XX 02-DEC-1999; 99WO-US28564.  
PR

(GETH ) GENENTECH INC.

PI Botstein D, Goddard A, Gurney AL, Hillan KJ, Roy MA, Wood WI;

DR PPI: 2001-091968/10.

XX P-PSDB; AAB68598.

PT New antibody that binds to a PRO polypeptide, e.g. PRO187 and PRO533,  
PT useful for diagnosing and treating cancers -

PS Claim 50; Fig 13; 196pp; English.

XX The present invention relates to PRO proteins and coding sequences. The  
CC present sequence is the coding sequence for one such PRO protein.  
CC It was found that the PRO genes are amplified in the genome of tumour  
CC cells. The gene amplification is expected to be associated with the  
CC overexpression of the gene product and contributes to tumorigenesis.  
CC Therefore, antagonists of PRO proteins are useful for the treatment of  
CC benign or malignant tumours, leukaemias, lymphoid malignancies and other  
CC disorders such as neuronal, glial, astrocytal, hypothalamic, glandular,  
CC epithelial, inflammatory and immunologic disorders.

XX Sequence 1266 BP; 216 A; 418 C; 390 G; 242 T; 0 other;

Query Match 5.0%; Score 32; DB 22; Length 1266;

Best Local Similarity 100.0%; Pred. No. 1.3e-05;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 375 GTACCCCTGCTGCTGATGCTGCTGCTG 406

Db 136 GTACCCCTGCTGCTGATGCTGCTGCTG 167

RESULT 14  
AAC97451  
ID AAC97451 standard; cDNA; 1266 BP.  
XX  
XX AAC97451;  
AC  
XX  
XX 28-FEB-2001 (first entry)  
DT  
XX  
XX Human angiogenesis-associated protein PRO261 cDNA, SEQ ID NO:107.  
DE  
XX  
XX Human; angiogenesis-associated protein; PRO; endothelial cell growth;  
XX cardiac hypertrophy; cardiovascular disorder; endothelial disorder;  
XX angiogenic disorder; atherosclerosis; osteoporosis; hypertension;  
XX myocardial infarction; diabetic retinopathy; rheumatoid arthritis;  
XX Crohn's disease; psoriasis; endometriosis; ulcer; wound healing; cancer;  
XX Alzheimer's disease; Huntington's disease; stroke; drug screening;  
XX gene therapy; transgenic animal; ss.

OS Homo sapiens.

XX WO200053753-A2.

XX 14-SEP-2000.

XX 05-JAN-2000; 2000WO-US00219.

XX 08-MAR-1999; 99WO-US05028.

XX 12-MAR-1999; 99US-0123957.

XX 14-MAY-1999; 99US-0134287.

XX 02-JUN-1999; 99WO-US12252.

XX 23-JUN-1999; 99US-0141037.

XX 20-JUL-1999; 99US-0144758.

XX 26-JUL-1999; 99US-0145698.

XX 01-SEP-1999; 99WO-US20111.

XX 08-SEP-1999; 99WO-US20594.

XX 15-SEP-1999; 99WO-US21547.

XX 05-OCT-1999; 99WO-US23089.

XX 30-NOV-1999; 99WO-US28313.

XX 02-DEC-1999; 99WO-US28564.

XX 02-DEC-1999; 99WO-US28565.

(GETH ) GENENTECH INC.

PI Ashkenazi AJ, Baker KP, Ferrara N, Gerber H, Goddard A;

PI Godowski PJ, Gurney AL, Hillan KJ, Kuo SS, Mark MR, Marsters SA;

PI Paoni NF, Pitti RM, Watanabe CK, Williams PM, Wood WI;

DR PPI: 2001-090793/10.

XX P-PSDB; AAB53084.

PT New isolated nucleic acid for producing a PRO polypeptide, analyzing  
PT genetic disorders and treating cardiovascular, endothelial or  
PT angiogenic disorders, such as atherosclerosis, wounds or cancer -

PS Claim 58; Fig 41; 293pp; English.

XX The invention relates to novel human angiogenesis-associated proteins  
CC designated PRO proteins (AAB53064-B53097), and to nucleic acids encoding  
CC PRO proteins. The invention also relates to vectors and host cells  
CC comprising a PRO nucleic acid, the recombinant production of a PRO  
CC protein, PRO antibodies specific for a PRO protein, fusion proteins  
CC comprising a PRO protein, agonists or antagonists of a PRO protein, and  
CC compounds which inhibit the expression of a PRO gene. The invention  
CC additionally encompasses methods of identifying modulators of PRO  
CC expression or activity; diagnosing a cardiovascular, endothelial or  
CC angiogenic disorder, or a susceptibility to such a disorder by detecting  
CC mutations in a PRO gene, or the expression level of a PRO gene within a



APPLICANT: Botstein, David A.  
APPLICANT: Cohen, Robert  
APPLICANT: Goddard, Audrey  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: Lawrence, David A.  
APPLICANT: Levine, Arnold J.  
APPLICANT: Pennica, Diane  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME  
FILE REFERENCE: P1176R2  
CURRENT APPLICATION NUMBER: US/09/182,145B  
EARLIER FILING DATE: 1998-10-29  
EARLIER APPLICATION NUMBER: US 60/063,704  
EARLIER FILING DATE: 1997-10-29  
EARLIER APPLICATION NUMBER: US 60/073,612  
EARLIER FILING DATE: 1998-02-04  
EARLIER APPLICATION NUMBER: US 60/081,695  
EARLIER FILING DATE: 1998-04-14  
NUMBER OF SEQ ID NOS: 156  
SEQ ID NO 18  
LENGTH: 1734  
TYPE: DNA  
ORGANISM: Mus musculus  
US-09-182-145-18

Query Match 14.2%: Score 90; DB 4; Length 1734;  
Best Local Similarity 100.0%; Pred. No. 4.2e-35;  
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 410 ACTGTGTGACGAGAGCTGGGGAGTCTGCGACCACTGCATGTCGTGGACCCAGCCA 469  
DB 1317 ACTGTGTGACGAGAGCTGGGGAGTCTGCGACCACTGCATGTCGTGGACCCAGCCA 1258  
QY 470 GGGCCTGCTTGTGACCTGGGGGACAGCCC 499  
DB 1257 GGGCCTGCTTGTGACCTGGGGGACAGCCC 1228

RESULT 3  
US-09-182-145-38  
Sequence 38, Application US/09182145B  
Patent No. 6387657  
GENERAL INFORMATION:  
APPLICANT: Botstein, David A.  
APPLICANT: Cohen, Robert  
APPLICANT: Goddard, Audrey  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: Lawrence, David A.  
APPLICANT: Levine, Arnold J.  
APPLICANT: Pennica, Diane  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME  
FILE REFERENCE: P1176R2  
CURRENT APPLICATION NUMBER: US/09/182,145B  
EARLIER FILING DATE: 1998-10-29  
EARLIER APPLICATION NUMBER: US 60/063,704  
EARLIER FILING DATE: 1997-10-29  
EARLIER APPLICATION NUMBER: US 60/073,612  
EARLIER FILING DATE: 1998-02-04  
EARLIER APPLICATION NUMBER: US 60/081,695  
EARLIER FILING DATE: 1998-04-14  
NUMBER OF SEQ ID NOS: 156  
SEQ ID NO 38  
LENGTH: 738  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-182-145-38

Query Match 5.0%: Score 32; DB 4; Length 738;

Best Local Similarity 100.0%; Pred. No. 1.6e-06;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 375 GTACCCCTGCTGATGAGCTGTGCTGCTG 406  
DB 115 GTACCCCTGCTGATGAGCTGTGCTGCTG 146

RESULT 4  
US-09-182-145-13  
Sequence 13, Application US/09182145B  
Patent No. 6387657  
GENERAL INFORMATION:  
APPLICANT: Botstein, David A.  
APPLICANT: Cohen, Robert  
APPLICANT: Goddard, Audrey  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: Lawrence, David A.  
APPLICANT: Levine, Arnold J.  
APPLICANT: Pennica, Diane  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME  
FILE REFERENCE: P1176R2  
CURRENT APPLICATION NUMBER: US/09/182,145B  
EARLIER FILING DATE: 1998-10-29  
EARLIER APPLICATION NUMBER: US 60/063,704  
EARLIER FILING DATE: 1997-10-29  
EARLIER APPLICATION NUMBER: US 60/073,612  
EARLIER FILING DATE: 1998-02-04  
EARLIER APPLICATION NUMBER: US 60/081,695  
EARLIER FILING DATE: 1998-04-14  
NUMBER OF SEQ ID NOS: 156  
SEQ ID NO 13  
LENGTH: 1293  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-182-145-13

Query Match 5.0%: Score 32; DB 4; Length 1293;  
Best Local Similarity 100.0%; Pred. No. 1.6e-06;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 375 GTACCCCTGCTGATGAGCTGTGCTGCTG 406  
DB 148 GTACCCCTGCTGATGAGCTGTGCTGCTG 179

RESULT 5  
US-09-182-145-14/C  
Sequence 14, Application US/09182145B  
Patent No. 6387657  
GENERAL INFORMATION:  
APPLICANT: Botstein, David A.  
APPLICANT: Cohen, Robert  
APPLICANT: Goddard, Audrey  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: Lawrence, David A.  
APPLICANT: Levine, Arnold J.  
APPLICANT: Pennica, Diane  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME  
FILE REFERENCE: P1176R2  
CURRENT APPLICATION NUMBER: US/09/182,145B  
EARLIER FILING DATE: 1998-10-29  
EARLIER APPLICATION NUMBER: US 60/063,704  
EARLIER FILING DATE: 1997-10-29  
EARLIER APPLICATION NUMBER: US 60/073,612  
EARLIER FILING DATE: 1998-02-04  
EARLIER APPLICATION NUMBER: US 60/081,695

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 28, 2003, 14:50:06 ; Search time 34.7003 seconds  
(without alignments)  
5612.050 Million cell updates/sec

Title: US-10-010-408-1\_COPY\_1\_635

Perfect score: 635  
Sequence: 1 GACGCTTCGTGATCTCCACAG.....GTGAGACGTGCTTCACCC 635

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 441362 seqs, 153338381 residues

Word size: 0

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database: Issued\_Patents\_NA.\*

1: /cgn2.6/ptodata/1/lna/5A.COMB.seq:\*  
2: /cgn2.6/ptodata/1/lna/5B.COMB.seq:\*  
3: /cgn2.6/ptodata/1/lna/6A.COMB.seq:\*  
4: /cgn2.6/ptodata/1/lna/6B.COMB.seq:\*  
5: /cgn2.6/ptodata/1/lna/PCBUS.COMB.seq:\*  
6: /cgn2.6/ptodata/1/lna/Backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Query length | ID                | Description       |
|------------|-------|-------------|--------------|-------------------|-------------------|
| 1          | 90    | 14.2        | 1734         | US-09-182-145-17  | Sequence 17, Appl |
| 2          | 90    | 14.2        | 1734         | US-09-182-145-18  | Sequence 18, Appl |
| 3          | 32    | 5.0         | 738          | US-09-182-145-38  | Sequence 38, Appl |
| 4          | 32    | 5.0         | 1293         | US-09-182-145-13  | Sequence 13, Appl |
| 5          | 32    | 5.0         | 1293         | US-09-182-145-14  | Sequence 14, Appl |
| 6          | 27    | 4.3         | 51           | US-09-182-145-117 | Sequence 117, App |
| 7          | 19    | 3.0         | 616          | US-09-385-982-220 | Sequence 220, App |
| 8          | 19    | 3.0         | 841          | US-09-182-145-38  | Sequence 39, Appl |
| 9          | 19    | 3.0         | 1196         | US-09-149-476-225 | Sequence 225, App |
| 10         | 19    | 3.0         | 1220         | US-09-149-476-57  | Sequence 57, Appl |
| 11         | 19    | 3.0         | 1514         | US-09-213-768-1   | Sequence 13, Appl |
| 12         | 13    | 3.0         | 1539         | US-09-668-680-13  | Sequence 13, Appl |
| 13         | 18    | 2.8         | 20           | US-09-213-768-2   | Sequence 2, Appl  |
| 14         | 18    | 2.8         | 2949         | US-08-253-451-10  | Sequence 10, Appl |
| 15         | 18    | 2.8         | 3727         | US-08-249-380-1   | Sequence 10, Appl |
| 16         | 18    | 2.8         | 8957         | US-08-259-451-1   | Sequence 152, App |
| 17         | 17    | 2.7         | 44           | US-09-182-145-152 | Sequence 206, App |
| 18         | 17    | 2.7         | 480          | US-09-188-930-206 | Sequence 151, App |
| 19         | 17    | 2.7         | 614          | US-08-998-416-151 | Sequence 3, Appl  |
| 20         | 17    | 2.7         | 661          | US-08-578-634C-3  | Sequence 3, Appl  |
| 21         | 17    | 2.7         | 896          | US-09-430-010-3   | Sequence 3, Appl  |
| 22         | 17    | 2.7         | 1491         | US-09-662-249A-3  | Sequence 36, Appl |
| 23         | 17    | 2.7         | 1491         | US-09-662-249A-3  | Sequence 20, Appl |
| 24         | 17    | 2.7         | 1743         | US-08-665-259-20  | Sequence 20, Appl |
| 25         | 17    | 2.7         | 1743         | US-08-762-500-20  | Sequence 78, Appl |
| 26         | 17    | 2.7         | 1974         | US-08-762-500-78  | Sequence 43, Appl |
| 27         | 17    | 2.7         | 3280         | US-09-484-970B-43 |                   |

|    |    |     |        |                   |                   |
|----|----|-----|--------|-------------------|-------------------|
| 28 | 17 | 2.7 | 3865   | US-09-149-476-296 | Sequence 296, App |
| 29 | 17 | 2.7 | 6803   | US-08-665-259-19  | Sequence 19, Appl |
| 30 | 17 | 2.7 | 6803   | US-08-762-500-19  | Sequence 19, Appl |
| 31 | 17 | 2.7 | 7874   | US-09-780-1175-96 | Sequence 96, Appl |
| 32 | 17 | 2.7 | 35828  | US-09-449-218D-17 | Sequence 17, Appl |
| 33 | 17 | 2.7 | 81001  | US-09-750-580-1   | Sequence 1, Appl  |
| 34 | 17 | 2.7 | 111282 | US-09-754-250-3   | Sequence 3, Appl  |
| 35 | 16 | 2.5 | 428    | US-09-397-787-224 | Sequence 224, App |
| 36 | 16 | 2.5 | 435    | US-09-397-787-280 | Sequence 280, App |
| 37 | 16 | 2.5 | 572    | US-09-343-653-5   | Sequence 5, Appl  |
| 38 | 16 | 2.5 | 691    | US-08-243-879A-37 | Sequence 37, Appl |
| 39 | 16 | 2.5 | 691    | US-08-499-523-3   | Sequence 3, Appl  |
| 40 | 16 | 2.5 | 691    | US-09-128-345-3   | Sequence 3, Appl  |
| 41 | 16 | 2.5 | 760    | US-08-768-859A-7  | Sequence 7, Appl  |
| 42 | 16 | 2.5 | 760    | US-08-767-820A-7  | Sequence 7, Appl  |
| 43 | 16 | 2.5 | 760    | PCR-US95-06157-7  | Sequence 7, Appl  |
| 44 | 16 | 2.5 | 766    | US-08-768-859A-9  | Sequence 9, Appl  |
| 45 | 16 | 2.5 | 766    | US-08-767-820A-9  | Sequence 9, Appl  |

## ALIGNMENTS

```
RESULT 1
US-09-182-145-17
; Sequence 17, Application US/09182145B
; Patent No. 6387657
; GENERAL INFORMATION:
; APPLICANT: Bolstein, David A.
; APPLICANT: Cohen, Robert
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Lawrence, David A.
; APPLICANT: Levine, Arnold J.
; APPLICANT: Pennica, Diane
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: WISP POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: P117682
; CURRENT APPLICATION NUMBER: US/09/182,145B
; CURRENT FILING DATE: 1998-10-29
; EARLIER APPLICATION NUMBER: US 60/063,704
; EARLIER FILING DATE: 1997-10-29
; EARLIER APPLICATION NUMBER: US 60/073,612
; EARLIER FILING DATE: 1998-02-04
; EARLIER APPLICATION NUMBER: US 60/081,695
; NUMBER OF SEQ ID NOS: 156
; SEQ ID NO 17
; LENGTH: 1734
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-182-145-17

Query Match          14.2%; Score 90; DB 4; Length 1734;
Best Local Similarity 100.0%; Pred. No. 4.2e-35;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 410 AGGTGTCACGAGAGCTGGGGAGTCTGCGACACCTGCATGTTCGACCCAGCCA 469  
DB 418 AGTGTGTCACGAGAGCTGGGGAGTCTGCGACACCTGCATGTTCGACCCAGCCA 477  
QY 470 GGGCTGGTGTGTCAGCTGGGGAGGCCC 499  
DB 478 GGGCTGGTGTGTCAGCTGGGGAGGCCC 507

RESULT 2  
US-09-182-145-18/C  
; Sequence 18, Application US/09182145B  
; Patent No. 6387657  
; GENERAL INFORMATION:

particular tissue; treating a cardiovascular, endothelial or angiogenic disorder via the administration of a PRO protein, PRO nucleic acid, or PRO agonist or antagonist; a retroviral gene therapy vector comprising a PRO nucleic acid; and methods of inhibiting or stimulating endothelial cell growth, cardiac hypertrophy or PRO-induced angiogenesis via the administration of a PRO protein, or an agonist or antagonist thereof. PRO nucleic acids, PRO proteins, antibodies against PRO proteins, PRO agonists and PRO antagonists may be used as therapeutic agents to treat cardiovascular, endothelial or angiogenic disorders, such as atherosclerosis, osteoporosis, myocardial infarction, hypertension, diabetic retinopathy, rheumatoid arthritis, Crohn's disease, psoriasis, endometriosis, ulcers, wounds, cancer, Alzheimer's disease, Huntington's disease, or stroke. PRO nucleic acids are additionally useful in the recombinant production of PRO proteins, as hybridisation probes to screen libraries to isolate cDNAs with sequence identity to PRO proteins, to map genes encoding PRO proteins, to analyse genetic disorders, and in gene therapy. PRO nucleic acids can also be used to produce transgenic animals useful for the development and screening of potential therapeutic agents. The present sequence represents a cDNA encoding a PRO protein of the invention.

XX Sequence 1266 BP; 216 A; 418 C; 390 G; 242 T; 0 other;

Query Match 5.0%; Score 32; DB 22; Length 1266;

Best Local Similarity 100.0%; Pred. No. 1.3e-05;

Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 375 GTACCCCTGGTGGATGGCTGTGGCTGCTG 406

DB 136 GTACCCCTGGTGGATGGCTGTGGCTGCTG 167

RESULT 15

AAA77566

ID AAA77566 standard; cDNA; 1267 BP.

XX

AC AAA77566;

XX

DT 07-NOV-2000 (first entry)

XX

DE Human PRO261 cDNA sequence SEQ ID NO:71.

XX

KW Human; PRO; promotion; inhibition; angiogenesis; cardiovascularisation; diagnosis; trauma; wound; cancer; atherosclerosis; cardiac hypertrophy; angiogenic; proliferative; cardiant; cardiovascular; antiatherosclerotic; cytostatic; gene therapy; vaccine; ss.

XX

OS Homo sapiens.

XX

PN WO200032221-A2.

XX

PD 08-JUN-2000.

XX

PF 30-NOV-1999; 99WO-US28313.

XX

PR 01-DEC-1998; 98WO-US25108.

PR 16-DEC-1998; 98US-0112850.

PR 12-JAN-1999; 99US-0115554.

PR 08-MAR-1999; 99WO-US05028.

PR 12-MAR-1999; 99US-0123957.

PR 28-APR-1999; 99US-0131445.

PR 14-MAY-1999; 99US-0134287.

PR 02-JUN-1999; 99WO-US12252.

PR 23-JUN-1999; 99US-0141037.

PR 20-JUL-1999; 99US-0144758.

PR 26-JUL-1999; 99US-0145698.

PR 01-SEP-1999; 99WO-US20111.

PR 08-SEP-1999; 99WO-US20594.

PR 13-SEP-1999; 99WO-US20944.

PR 15-SEP-1999; 99WO-US21090.

PR 15-SEP-1999; 99WO-US21547.

PR 05-OCT-1999; 99WO-US23089.

PR 29-OCT-1999; 99US-0162506.

XX (GETH ) GENENTECH INC.

XX

PA Ashkenazi AJ, Baker KP, Ferrara N, Gerber H, Hillan KJ, Goddard A;

PI Godowski PJ, Gurney AL, Klein RD, Kuo SS, Paoni NF, Smith V;

PI Watanabe CK, Williams PM, Wood WI;

XX

XX WPI; 2000-412154/35.

DR

DR P-PSDB; AAB24402.

XX

XX Nucleic acids encoding PRO polypeptides useful for preventing, diagnosing and treating disorders in mammals -

PT

PT Claim 61; Fig 29; 315pp; English.

XX

CC The present invention describes nucleic acids encoding PRO polypeptides useful for preventing, diagnosing and treating disorders in mammals by cardiovascular, endothelial or angiogenic disorder in mammals by modulating cell proliferation, angiogenesis and cardiovascularisation, and for identifying agonists and antagonists of these processes. The nucleic acids and the proteins they encode may be used in the prevention, treatment and diagnosis of diseases associated with inappropriate PRO expression such as cardiovascular, endothelial or angiogenic disorders in mammals (e.g. atherosclerosis, cancers and cardiac hypertrophy). For example, the nucleic acids (NCs) and vectors containing them and the PRO polypeptide may be used to treat disorders associated with decreased PRO expression. AAA77510 to AAA77721 and AAB24388 to AAB24435 represent nucleotide and protein sequences used in the exemplification of the present invention.

XX

SQ Sequence 1267 BP; 216 A; 419 C; 390 G; 242 T; 0 other;

Query Match

Best Local Similarity 5.0%; Score 32; DB 21; Length 1267;

Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 375 GTACCCCTGGTGGATGGCTGTGGCTGCTG 406

DB 136 GTACCCCTGGTGGATGGCTGTGGCTGCTG 167

Search completed: July 28, 2003, 15:54:57

Job time : 167.775 secs

THIS PAGE BLANK (USPTO)